



Functional Brain Imaging by EEG: A Window to the Human Mind

Stahlhut, Carsten

Publication date:
2012

Document Version
Publisher's PDF, also known as Version of record

[Link back to DTU Orbit](#)

Citation (APA):
Stahlhut, C. (2012). *Functional Brain Imaging by EEG: A Window to the Human Mind*. Technical University of Denmark. IMM-PHD-2012 No. 254

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Functional Brain Imaging by EEG: A Window to the Human Mind

Carsten Stahlhut

Kongens Lyngby 2011
IMM-PHD-2011-XXXXX

Technical University of Denmark
Department of Informatics and Mathematical Modelling
Building 321, DK-2800 Kongens Lyngby, Denmark
Phone +45 45253351, Fax +45 45882673
reception@imm.dtu.dk
www.imm.dtu.dk

IMM-PHD: ISSN 0909-3192

Abstract

This thesis presents electroencephalography (EEG) brain imaging by covering topics as empirical evaluation of source confusion, probabilistic inverse methods, and source analysis performed on infant EEG data. In terms of source confusion we inspect how current sources within the brain may be confused with each other as noise is present in the EEG recordings. Moreover, we examine how errors in the forward model affect the source confusion.

The primary aim of this thesis is to provide sharper EEG brain images by improving current inverse methods. In this relation we focus the attention on two topics in EEG source reconstruction, namely, the forward propagation model (describing the mapping from the current sources within the brain to the sensors at the scalp) and the temporal patterns present in the EEG.

As forward models may suffer from a number of errors including the geometrical representation of the human head, the tissue conductivity distribution, and electrode positions, we propose an algorithm which consider forward model uncertainties. Bayesian graphical models provide a powerful means of incorporating prior assumptions that narrow the solution space and lead to tractable posterior distributions over the unknown sources given the observed data. Here, we propose a hierarchical Bayesian model that attempts to minimize the influence of uncertainties associated with the forward model on the source estimates.

Similarly, we develop a hierarchical spatio-temporal Bayesian model that accommodates the principled computation of sparse spatial and smooth temporal EEG source reconstructions consistent with neurophysiological assumptions in a variety of event-related imaging paradigms.

Resumé (Abstract in Danish)

Denne afhandling præsenterer hjernebilleddannelse ved brug af elektroencephalografi (EEG), hvor emner såsom empirisk evaluering af kildeforvirring, probabilistiske inverse metoder, og kildeanalyse hos spædbørn behandles. I forbindelse med kildeforvirring undersøges det hvordan kilder i hjernen risikerer at blive forvekslet med andre kilder i hjernen som følge af støj i EEG. Endvidere undersøges hvilken indflydelse fejl i hovedmodellen har på omfanget af kildeforvirring.

Afhandlingens primære formål er at levere skarpere EEG hjernebilleder i form af udvikling af forbedrede inverse metoder. I den sammenhæng fokuses der på to hovedelementer for EEG kilderekonstruktion: Hovedmodellen (forbindelsesledet mellem kilderne i hjernen og sensorerne placeret på hovedbunden) samt de temporale mønstre i EEG signalet.

Da der er risiko for at hovedmodellerne kan være behæftet med forskellige typer af fejl, herunder den geometriske repræsentation af hovedet, fordelingen af ledningsevne samt elektrode koordinaterne, foreslås der en algoritme, som tager højde for usikkerheder i hovedmodellen. Bayesianiske grafiske modeller er et effektivt redskab til at inkludere a-priori antagelser, således at udfaldsrummet af mulige løsninger begrænses og der herved kan opnås håndterbare posterior fordelinger af de ukendte kilder givet et observeret datasæt. Vi introducerer her en hierarkisk Bayesianisk model, der forsøger at minimere indflydelsen af usikkerheder relateret til hovedmodellen på kildeestimerne.

Ligeledes præsenteres en hierarkisk spatio-temporal Bayesianisk model, der tillader EEG kilderekonstruktion bestående af få spatiale kilder med tilhørende glatte

tidsfunktioner. Dette er i overensstemmelse med neurofysiologiske antagelser i en del event-relaterede paradigmer.

Preface

This thesis has been prepared in the section for Cognitive Systems at the Department of Informatics and Mathematical Modelling, Technical University of Denmark (DTU) in partial fulfillment of the requirements for acquiring the PhD degree in engineering.

The thesis consists of a summary report, a research abstract, and a collection of four research papers written during the period 2008-2011, and elsewhere published.

Lyngby, April 2011

Carsten Stahlhut

Publications included

Papers

- [A] Carsten Stahlhut, Morten Mørup, Ole Winther, and Lars K. Hansen. SOFOMORE: Combined EEG Source and Forward Model REconstruction. In *Biomedical Imaging: From Nano to Macro, 2009, ISBI'09. IEEE International Symposium on*, 2009, pages 450-453.
- [B] Carsten Stahlhut, Morten Mørup, Ole Winther, and Lars K. Hansen. Hierarchical Bayesian Model for simultaneous EEG Source and Forward Model Reconstruction (SOFOMORE). In *Machine Learning for Signal Processing (MLSP), MLSP'09, IEEE International Workshop on*, 2009, pages 1-6.
- [C] Carsten Stahlhut, Morten Mørup, Ole Winther, and Lars K. Hansen. Simultaneous EEG Source and Forward Model Reconstruction (SOFOMORE) using a Hierarchical Bayesian Approach. In *Journal of Signal Processing System (JSPS)*. 2010, pages 1-14.
- [D] Carsten Stahlhut, Hagai T. Attias, David Wipf, Lars K. Hansen, and Srikantan S Nagarajan. Sparse Spatio-temporal Inference of Electromagnetic Brain Sources. In *Machine Learning in Medical Imaging*, volume 6357 of *Lecture Notes in Computer Science* 2010, pages 157-164.

Abstract

- [E] Carsten Stahlhut, Morten Mørup, Ole Winther, and Lars K. Hansen. Evaluation of the Influence of uncertain Forward Models on the EEG Source Reconstruction Problem. The 15th Annual Human Brain Mapping Meeting. June 2009.

Acknowledgements

First of all I wish to express my deepest gratitude to my two supervisors Professor Lars Kai Hansen and Associate Professor Ole Winther for providing continuous support and guidance. I am truly happy that you both opened my eyes to this exciting world of neuroscience and machine learning years ago. And maybe even more importantly for always being filled with enthusiasm even if you had to explain/discuss your ideas several times.

A special thanks to Assistant Professor Morten Mørup, DTU Informatics, for our collaboration, all your exciting ideas and highly valuable feedback. I also wish to thank Associate Professor Jan Larsen and Tobias Andersens, DTU Informatics, for sharing your knowledge about machine learning and cognitive processes. Similarly, I want to thank the rest of my colleagues in Cognitive Systems for the social and professional interactions that we have had.

I wish to thank Associate Professor Sid Kouider (Ecole Normale Supérieure) as well, for introducing me to the subfield of neuroscience in infants. I am very happy for our collaboration and your enthusiasm in the field.

During my studies I spent 5 months at University of California, San Francisco (UCSF), which I enjoyed very much. Of importance for making the visit so successful and inspiring is Professor Srikantan S. Nagarajan. I am very grateful that you gave me the opportunity to visit your lab, sharing many of your interesting ideas, and for all the fruitful discussions we had. Likewise, I would like to thank David Wipf (UCSF) and Hagai T. Attias (Convex Brain Imaging) for our interesting discussions and allowing me to benefit from your great experience with machine learning methods. And of course not to forget the rest of the

Biomagnetic Imaging Laboratory (BIL) and especially Julia P. Owen and the people involved in the monthly BIL sushi group. Undoubtly, you all made my visit at UCSF unforgettable. I hope our paths will cross again in the nearest future.

Additionally, I want to thank Morten Hansen and Michael Kai Petersen for proofreading this thesis.

Finally, I wish to thank my wife Michelle Stahlhut for your continuous encouragement and support. And even more importantly for our lovely daughter Sofia and for making the visit at UCSF possible by accompanying me on the journey as a family.

Abbreviations

aFDM	anisotropic Finite Difference Method
ARD	Automatic Relevance Determination
BEM	Boundary Element Method
BIC	Bayesian Information Criteria
CS	Cheeseman-Stutz
CT	Computer Tomography
DTI	Diffusion Tensor Imaging
ECD	Equivalent Current Dipoles
EEG	Electroencephalography
EP	Expectation Propagation
ERP	Event Related Potential
ESP	Event Sparse Penalty
FDM	Finite Difference Method
FEM	Finite Element Method
fMRI	functional Magnetic Resonance Imaging
FP	False Positive
ICA	Independent Component Analysis
iFDM	isotropic Finite Difference Method
KL	Kullback-Leibler

MAP	Maximum A-Posterior
MCMC	Markov Chain Monte Carlo
MEG	Magnetoencephalography
MN	Minimum Norm
MRI	Magnetic Resonance Imaging
NIRS	Near InfraRed Spectroscopy
PCA	Principle Component Analysis
PET	Positron Emission Tomography
PPV	Positive Prediction Value
ReML	Restricted Maximum Likelihood
sls	Sparse least squares
SNR	Signal-to-Noise Ratio
SOFOMORE	SOURCE and FORWARD MODEL REconstruction
SPECT	Single Photon Emission Computed Tomography
STE	Space-Time Events
TLS	Total Least Squares
TP	True Positive
VB	Variational Bayes

Notation

General terms

x	Scalar
\mathbf{x}	Column vector
x_i	The i th element of \mathbf{x}
$\mathbf{x}_{\setminus i}$	Excluding the i th element in \mathbf{x}
\mathbf{x}_i	The i th "vector element" of \mathbf{x}
$\mathbf{x}_{1:N}$	The vector containing $[\mathbf{x}_1^T, \dots, \mathbf{x}_N^T]^T$
\mathbf{X}	Matrix
$(\mathbf{X})_{i,j}$	The (i,j) th element of matrix \mathbf{X}
$X_{i,j}$	The (i,j) th element of matrix \mathbf{X}
\mathbf{I}_M	Identity matrix of size $M \times M$
$\langle \cdot \rangle$	Statistical expectation operator
$ \cdot $	Absolute value
$p(\cdot)$	Probability distribution
$\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$	Gaussian distribution with mean $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$
$\mathcal{G}(\nu, \zeta)$	Gamma distribution with shape ν and rate ζ parameters
$\psi(\cdot)$	The digamma function

Set operators

\mathbb{R}	The set of real numbers
\in	Belongs to

Vector operators

$\text{diag}(\mathbf{x})$	Diagonal matrix with \mathbf{x} in the diagonal
---------------------------	---

Matrix operators

$(\cdot)^{-1}$	Inverse matrix
$(\cdot)^T$	Transposed matrix
\mathbf{X}^+	Pseudo inverse matrix of the matrix \mathbf{X}
\mathbf{X}^n	The n th power of a square matrix \mathbf{X}
$\text{diag}(\cdot)$	Vector given by diagonal of the matrix
$\det(\cdot)$	Determinant of matrix
$\text{tr}(\cdot)$	Trace operation
$\sigma_i(\mathbf{X})$	The i th singular value of \mathbf{X}

Variable names

N_c	Number of channels/sensors
N_d	Number of dipoles/sources
N_t	Number of time points/samples
N_k	Number of basis functions
\mathbf{M}	Measurements of size $N_c \times N_t$
\mathbf{A}	Forward model of size $N_c \times N_d$
\mathbf{S}	Sources of size $N_d \times N_t$
\mathbf{W}	Spatio-temporal maps of size $N_k \times N_d$
Φ	Temporal basis functions of size $N_k \times N_t$
θ	The set of variables in a model

Contents

Abstract	i
Resumé (Abstract in Danish)	iii
Preface	v
Publications included	vii
Acknowledgements	ix
Abbreviations	xi
Notation	xiii
1 Introduction and Motivation	1
1.1 Thesis Outline and Contributions	2
2 Preliminaries	5
2.1 The Human Brain	5
2.2 The Forward Problem	7
2.3 Inverse Problem	13
3 Source Confusion	19
3.1 Qualitative Analysis of the Role of Forward Model Uncertainty .	19
3.2 Related Work	24
3.3 Summary	26
4 Source Localization in Forward Model Uncertainty	29
4.1 Method: The SOFOMORE Model	30
4.2 Empirical Evaluation	34

4.3	Related Work	35
4.4	Summary	36
5	Spatio-Temporal Approach for Source Localization	41
5.1	Method: The Aquavit Model	42
5.2	Empirical Evaluation	45
5.3	Related Work	48
5.4	Summary	50
6	Source Analysis in Infants	53
6.1	Paradigm	54
6.2	Method	55
6.3	Empirical Evaluation	56
6.4	Related Work	61
6.5	Summary	62
7	Conclusion	73
A	SOFOMORE: Combined EEG SOurce and FOrward MOdel REconstruction	77
B	Hierarchical Bayesian Model for simultaneous EEG Source and Forward Model Reconstruction (SOFOMORE)	83
C	Simultaneous EEG Source and Forward Model Reconstruction (SOFOMORE) using a Hierarchical Bayesian Approach	91
D	Sparse Spatio-temporal Inference of Electromagnetic Brain Sources	105
E	Evaluation of the Influence of uncertain Forward Models on the EEG Source Reconstruction Problem	115
	Bibliography	119

CHAPTER 1

Introduction and Motivation

Throughout centuries humans have been dedicated to expand their existing knowledge by addressing new and unsolved topics that surround us in our everyday life. Among one of these unsolved topics we find the complex structured human brain and its many functionalities, which is far from being fully understood. Here, neuroscience has played a key role in an attempt to crack the human brain code by exploring the anatomical structure and how structural pathways link different brain regions together. Either by structural neuroimaging modalities or by examining the functionality behavior of different brain regions through functional imaging. Among others, the Computer Tomography (CT), Magnetic Resonance Imaging (MRI), and Diffusion Tensor Imaging (DTI) can be mentioned in the category structural imaging modalities. Of functional imaging modalities we have e.g. Positron Emission Tomography (PET), functional MRI (fMRI), Near InfraRed Spectroscopy (NIRS), and magneto-/electroencephalography (M/EEG).

In terms of high spatial resolution in functional imaging modalities especially fMRI has made profound contributions to neuroscience despite of its poor temporal resolution (2s for whole brain imaging). In contrast, EEG provides very good temporal resolution (below 10ms), but in the current state of art it lacks spatial specificity. However, advanced statistical modeling methods are changing this picture and a better understanding of the spatial distribution of EEG sources and thus brain information processing has allowed the community to

produce sharper brain images of cortical activation in realistic experimental conditions.

This brings us directly to the focus of this PhD thesis, which is to pursue the goal of contributing to a better understanding of the EEG source distribution by examining the spatial distribution of the EEG sources and moreover to develop EEG brain imaging methods for producing sharper brain images. The challenges to sharp EEG brain imaging derive from the fact that the electrostatic inverse problem (mapping from surface electrodes to cortical sources) is extremely ill-posed, and that the forward propagation (from cortical sources to surface electrodes) is only partially observable.

1.1 Thesis Outline and Contributions

Chapter 2, Preliminaries, gives an introduction to neural activity and electroencephalography (EEG). The forward problem is treated and in this relation a description of the different types of head models that are used for EEG source localization is given. Additionally, the relationship between the forward and inverse problem is provided followed by a short subsection concerning Bayesian inference for source estimation.

Chapter 3, Source Confusion, presents an analysis of the forward model uncertainty. In this chapter it will appear how likely source activity arising from different locations in the brain may be confused with other locations in the brain.

Chapter 4, Source Localization in Forward Model Uncertainty, presents a method for performing combined Source and Forward Model REconstruction (SOFOMORE). The novelty of the proposed method is in its ability to perform forward model correction simultaneously with source estimation. The reasoning behind this approach is to minimize the influence of potential errors in the forward model on the source estimates.

Chapter 5, Spatio-Temporal Approach for Source Localizaiton, deals with the important spatio-temporal properties which exist in EEG. More specifically, a sparse spatio-temporal approach to solve the source reconstruction problem is presented in this chapter. The novelty of this method is in the sparsity in space and in the number of basis functions used to describe the temporal patterns that might exist in the given EEG of interest. Additionally, this method is based on the concept of calculating posterior distributions over the sources in stead of only point estimates. The advantage of posterior distributions compared to point estimates is that these also take the discrepancy of the data from its mean value into consideration.

Chapter 6, Source Analysis in Infants, introduces a new promising subfield in the neuroimaging society namely functional activity in infants. In this chapter the focus will be on how infants process and decode visual information such as faces. The analysis applies the standard source reconstruction approaches minimum norm and low resolution electromagnetic tomography [Hämäläinen and Ilmoniemi, 1994, Pascual-Marqui et al., 1994]. In order to minimize potential errors in the forward model the log evidence of a model is proposed for performing model selection between a number of different forward models.

Chapter 7, Conclusion, summarizes the thesis and proposes a selection of interesting future research directions.

Contributions

Besides the published work highlighted below, this thesis contributes with an extended analysis (in Chapter 3) of the work originally presented in Abstract E concerning how corrupted forward models may deteriorate the source estimates. Additionally, the source analysis in infants given in Chapter 6 is currently in preparation. The main contribution for source analysis in infants is an attempt to perform model selection between a set of forward models. Description of the published work is ordered according to the structuring of the chapters in this thesis.

Abstract E, Evaluation of the Influence of uncertain Forward Models on the EEG Source Reconstruction Problem, examines how noise affects reconstruction of sources located at different positions in the brain given different assumptions about the head model.

Paper A, SOFOMORE: Combined EEG Source and Forward Model REconstruction, contains a first attempt for estimating the current sources within the brain while jointly optimizing the forward model.

Paper B, Hierarchical Bayesian Model for simultaneous EEG Source and Forward Model Reconstruction (SOFOMORE), proves that the simultaneous source and forward model reconstruction is competitive to a Bayesian formulation of the minimum norm approach. Additionally, it is demonstrated that SOFOMORE is competitive to a similar hierarchical model as the SOFOMORE but just without forward model correction.

Paper C, Simultaneous EEG Source and Forward Model Reconstruction (SOFOMORE) using a Hierarchical Bayesian Approach, is an

extended comparison of the inverse methods in Paper [B](#). We explore the performance of the methods source estimates at different noise levels. Additionally, we examine what changes are made to the forward model. Finally, it includes detailed derivations of the SOFOMORE model.

Paper [D](#), Sparse Spatio-temporal Inference of Electromagnetic Brain Sources, proposes a probabilistic method for describing electromagnetic current sources as sparse in space and smooth in time. This is performed by the use of ARD priors on the sources in space as well as on temporal basis functions. The model reveals promising properties of being capable of balancing the source estimates on the temporal smoothness given by the temporal basis functions when these are consistent with the temporal patterns observed in the EEG signal. When the underlying temporal patterns do not have a support in the temporal basis functions the model shift its attention to a second term assuming independence across time such that it can capture source activation outside the span of the temporal basis set.

CHAPTER 2

Preliminaries

In this section fundamental theory related to Electroencephalography (EEG) will be introduced in order to understand how the EEG signal is generated and what realistic modeling assumptions can be made for the EEG inverse problem. The relation between the EEG forward and inverse problem is given in subsection 2.2 and 2.3, respectively.

EEG is actually a quite old measurement technique which was invented by the German physician Hans Berger in 1924 where he measured traces of electrical brain activity in humans. In principle, the foundation of the EEG technique was already completed in the late 18th century by Galvani, [Baillet et al., 2001]. He found electric activity in animals and named it "animal electricity". Today it is better known as electrophysiology.

2.1 The Human Brain

The human brain is a highly complex network consisting of $10^{10} - 10^{11}$ neurons and many more glial cells constituting the main 'building blocks' of the brain. According to [Phillips, 2001], "*the glial cells ensure the physical structure of the brain, the proper concentration of ions and the transport of the nutrients*

between blood vessels and brain tissue". The neurons on the other hand act as the information processing units of the brain. The structure of a neuron is illustrated in Fig. 2.1. A neuron can be divided into a cell body (also known as the soma), dendrites, and axons.

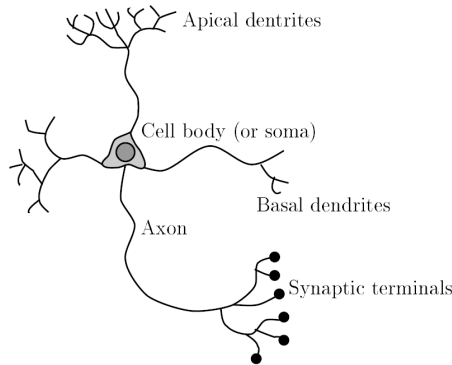


Figure 2.1: **A neuron** and its compartments - figure from [Phillips, 2001].

When a neuron is excited by others the apical dendrite membrane becomes transiently depolarized with respect to the soma and basal dendrites, as excitatory postsynaptic potentials (EPSP) are generated at the apical dendrites, [Baillet et al., 2001].

As a consequence of the depolarization and the potential difference between the apical dendrites versus soma and basal dendrites current flows from the apical dendrite to the soma and basal dendrites. Some parts of the current are intracellular and other parts are extracellular current loops starting with a current flowing from the extracellular surroundings through the neuron and back again to the extracellular surroundings. Intracellular current is normally denoted primary current, whereas the extracellular current is denoted the secondary, return, or volume currents. With the law of conservation of electric charges even the most distant point of the volume conductor leads to contribution to the extracellular current.

A closer look at the organization of the neurons reveals important model reduction information to the inverse problem (see Section 2.2). The different elements of the neurons can be found in different layers of the brain. In a thin outer layer of the cortex named the gray matter, the cell bodies and dendrites appear, while the axons are found in an interior layer called the white matter. From a macroscopic point of view it is believed that it is the neurons and the orientation of their apical dendrites that have an important influence on the ability to record the neural activity by EEG and MEG. In fact, it is believed that it is the

current (both primary and secondary) associated with the pyramidal cells (a special class of neurons) that give rise to the measured EEG or MEG signal, as their apical dendrites are organized parallel to each other and passing through different layers in the gray matter. Given the parallel structure of the apical dendrites and current passing through them the orientation of the current will be perpendicular to the cortical sheet. This is an important observation when dealing with the EEG inverse problem and modeling assumptions. However, the signal from a single neuron is far too weak to be detected by scalp recordings in EEG and it is thus believed that the main contribution instead is based on tens of thousands of neurons being active synchronously. Thus, for modeling purposes we may represent signal arising from a bundle of neurons by one common dipolar source instead of modeling the signal contribution from each single neuron. Combining tens of thousands of neurons in a bundle reduces the inverse problem significantly compared to solving an inverse problem with neurons as the spatial resolution.

2.2 The Forward Problem

As we are interested in the reconstruction of the source generators (within the brain) of the electrical fields measured in EEG by an array of sensors placed on the scalp, a relationship between the sensors and the sources needs to be established. Reconstruction of the EEG source generators is known as a severe ill-posed inverse problem with a mapping from source space to sensor space that is many to one. Despite the ill-posed nature of EEG source reconstruction it represents a key problem of high interest for the functional imaging community due to the EEG signal's high temporal resolution and relatively direct coupling to the neural signal, compared to, e.g., functional magnetic resonance imaging (fMRI) and positron emission tomography (PET).

Determining the relationship between sensors and sources, is known as solving the forward problem with a mapping from sources to the sensors. To solve the forward problem we first need to draw the attention towards the basis of EEG. According to [Plonsey and Heppner, 1967], no charge can be stored in the conducting extracellular volume for the signals of interest in the EEG, since the measured neural activity changes are slow compared to the propagation effects, [Hallesz et al., 2007]. Thus, the system does not introduce any time delays meaning that all fields at one point in time are generated by the active electric sources. Such a condition is denoted as being quasi-static as all fields and currents appear as being stationary at each time point yet with the possibility of changing over time. Under the quasi-static approximation to Maxwell's equations the electric fields and magnetic fields are decoupled in the sense that

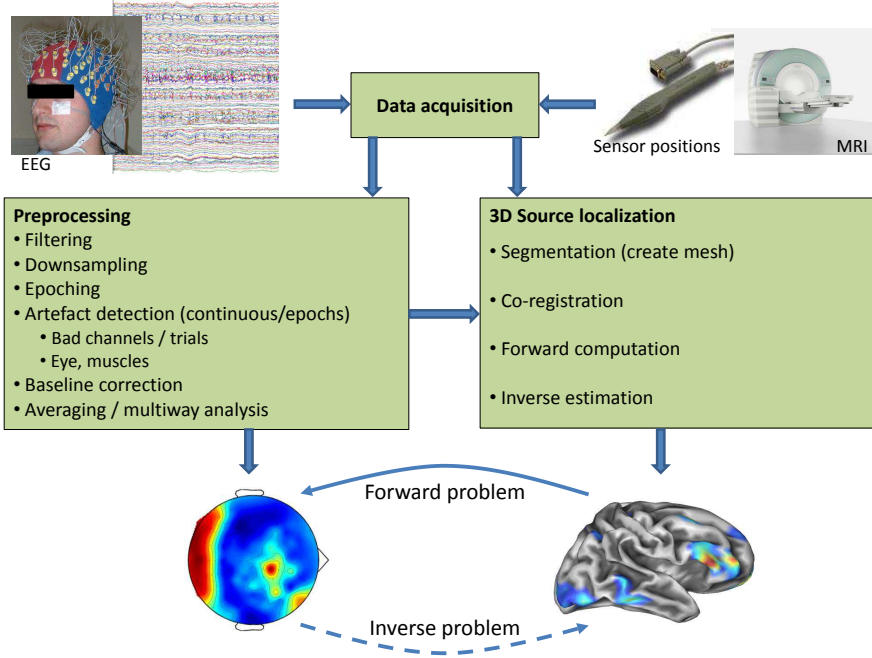


Figure 2.2: **Work flow for source localization.** Data acquisition includes EEG but can also include e.g. sensor positions, individual MRI scans, or other imaging modalities. In order to minimize the influence of noise in the EEG measurements preprocessing of the data should to be performed. A number of preprocessing steps can be done as illustrated in the Preprocessing box. Given the preprocessed EEG data and a forward model constructed by segmenting a MRI image (individual/template), co-registrate sensor space to MRI space, and solving the forward problem, the source localization can be performed. The relation between the forward and inverse problem is shown in the bottom of the figure.

the electric fields can be calculated or measured as if the magnetic fields do not exist and vice versa [Nunez and Srinivasan, 2006]. Thus, the relation between the measured EEG signal and the brain's current sources can be expressed as a linear instantaneous form in the sources. In particular, if the measured EEG signal is denoted $\mathbf{M} \in \mathbb{R}^{N_c \times N_t}$ and the current sources $\mathbf{S} \in \mathbb{R}^{N_d \times N_t}$, with N_c , N_d , and N_t being the number of channels, dipoles, and time samples, respectively, the forward relation can be written, [Baillet et al., 2001],

$$\mathbf{M} = \mathbf{A}\mathbf{S} + \mathcal{E}. \quad (2.1)$$

Here the noise \mathcal{E} is assumed additive. The interrelationship between the sensors and the current sources is given by the lead field matrix/forward model $\mathbf{A} \in \mathbb{R}^{N_c \times N_d}$ with the rows referred to as the lead fields for the sensors and the columns as the forward fields for the sources. With the prior knowledge of the orientation of the pyramidal cells from the previous section in mind, we can now restrict the orientation of the dipoles to be perpendicular to the cortex. Throughout the rest of this thesis we will apply this belief. However, if a more flexible orientation is desired e.g. if we want to reconstruct sources in volumen instead of on a surface, a flexible orientation can easily be incorporated by inclusion of three columns in \mathbf{A} per dipole corresponding to the three directional components. In fact, representing a dipole with three directional components can also be regarded as correcting the associated forward fields to some extent as modifying the orientation would be similar to changing the surface of the cortex such that its normal corresponds to the orientation of the dipole.

For estimating the forward fields of the dipoles a number of requirements are needed. This includes a model of the head geometry, sensor positions, as well as tissue conductivity values for the tissues involved in the head model. Obviously, with errors affecting the fundamental elements when computing the forward fields, this will propagate directly to the source estimates and affect these as well.

A schematic overview of the steps involved when performing 3D source reconstruction is found in Fig. 2.2. Moreover, the relation between the forward and inverse problem is shown.

2.2.1 Head Models

In order to describe the forward fields associated with current sources within the brain a volume conductor model of the head describing the varying conductivity layers in the head is needed. Modeling the conductivity layers is a necessity as the layers (tissues) attenuate the EEG signals differently due to their varying conductivity values. Different levels of complexity of the head model exist with the head models normally divided into spherical head models (unrealistic geometry) and realistic head models. Readers interested in a more detailed review of the approaches for solving the forward problem in EEG source analysis can e.g. consult [Mosher et al., 1999, Halletz et al., 2007].

Spherical Head Models

The most simple and first volume conductor model applied is a single homogeneous sphere, neglecting that the skull conductivity is significantly lower than the conductivity of the brain tissue and the scalp. A natural refinement of the

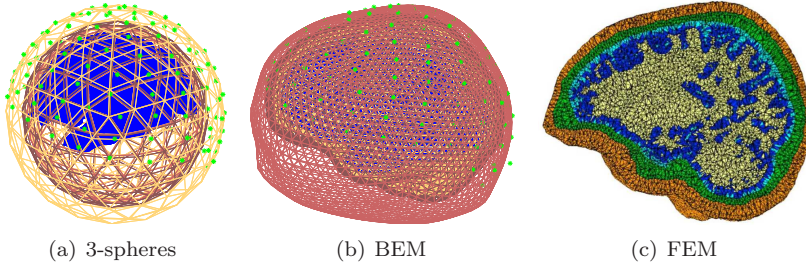


Figure 2.3: Illustration of **head models with increasing complexity**: 3-spheres head model, Boundary Element Method (BEM), and Finite Element Method. The 3-spheres and BEM were generated in the SPM8 academic software using routines from the Fieldtrip toolbox. The FEM is from [Wolters et al., 2006].

single sphere model is the 3-spheres head model consisting of three nested shells used to represent the brain, skull, and scalp. Figure 2.3(a) shows an example of a 3-spheres head model, where these layers are represented.

Realistic Head Models

The 3-spheres head model is a very simplified model of a human head. In reality heads are anisotropic, inhomogeneous, and not spherical. Especially, the skull and white matter tissue is considered highly anisotropic. According to [Hallez et al., 2007] the tangential conductivity is 10 times the conductivity perpendicular to the surface. Similarly, white matter tissues is regarded as anisotropic as it mainly consists of groups of axons (nerve bundles) and the conductivity along a nerve bundle is 9 times higher than perpendicular to it, [Hallez et al., 2007]. Thus, more realistic head models are needed, to obtain a more accurate solution to the forward problem. These head models use high resolution volumetric brain images, e.g. MRI, to extract anatomical information. Of realistic head models the Boundary Element Method (BEM), Finite Element Method, Finite Difference Method (FDM) can be mentioned.

Boundary Element Method: In a BEM model a piecewise homogeneous volume conductor is assumed and the surface potentials triggered by current sources are calculated numerically. Even though this model applies isotropic conductivities for the different tissues involved, it is still a widely used technique due to its relatively low computational cost and the realistic head geometry. As the name implies, a BEM model consists of surfaces enclosing a particular tissue. In practice each surface is tessellated with small boundary elements - triangles. Of most widely used BEM models is head models constructed by three surfaces: brain-skull, skull-scalp, and scalp-air interface, [Hallez et al., 2007, Fuchs et al.,

2002, Phillips, 2001, Ataseven et al., 2008]. An example of a BEM head model is found in Fig. 2.3(b).

Finite Element Method: An alternative approach for constructing a realistic head model is the finite element method (FEM), [Wolters et al., 2002, Wolters et al., 2006, Wolters et al., 2007]. This approach digitizes the entire 3D volume conductor into small elements typically tetrahedrons, which is in contrast to the BEM, where it is the surfaces that are tessellated with small elements. The most appealing factor of the FEM is the possibility of modeling tissues as anisotropic. A drawback of the FEM is, that it is normally regarded as being quite time consuming, due to the greatly increased computation complexity involved by tessellating a volumen in stead of a surface as in the BEM. However, significant improvements of the computational speed have been demonstrated by [Wolters et al., 2004, Steinsträter et al., 2010] using the open source SimBio software environment (<https://www.mrt.uni-jena.de/simbio>).

Finite Difference Method: An alternative realistic head model can be obtained by the finite difference method, which utilizes a cubic grid, [Saleheen and Ng, 1997, Hallez et al., 2005]. In fact, the FDM can be divided into two subgroups: the isotropic FDM (iFDM) and the anisotropic FDM (aFDM). As their names indicate they incorporate different assumption of a particular tissue being either isotropic or anisotropic, in the iFDM and aFDM, respectively.

In summary, the difference between BEM, FEM, and the FDM is the domain in which they are calculated. The BEM solves the forward problem by boundaries between homogeneous isotropic compartments, whereas the FEM and FDM solves the forward problem with an entire volume. Consequently, the FEM and FDM easily lead to larger number of computational points than the BEM. Typically, values for a BEM is in the order of 5,000 to 25,000 computational points whereas the computational points involved in FEM/FDM solutions are in the range 500,000 to 1,000,000. In the FDM the computational points are restricted to be cubic form, which can both be regarded as an advantage and disadvantage. It is an advantage in the sense that it is very easy to generate a structured FDM grid from segmented medical images as they are likewise based on cubic voxels. However, a drawback with the cubic structure compared to the grid used by FEM models is that it is not as flexible as the FEM in which the computational points (vertices of the tetrahedrons) can be choosen freely, [Hallez et al., 2007]. A direct result of this is a better representation of irregular interfaces between the different compartments using a FEM head model.

As the work presented in this thesis is mainly carried out using two academic software tools the SPM8 (<http://www.fil.ion.ucl.ac.uk/spm/>) and FieldTrip (<http://fieldtrip.fcdonders.nl>) the choices of forward models are limited to

spherical head models and BEM models. Unfortunately, these software packages do not support solutions to the forward problem using the FEM or FDM at the moment.

Given a solution to the forward problem, \mathbf{A} , the inverse problem can be approached. However, before we address the inverse problem let us examine the forward field matrix and the numerical issues related to it. As highlighted a number of times already EEG source reconstruction is an ill-posed problem. An ill-posed problem is the antagonist to a well-posed problem, which by definition of the french mathematician J. S. Hadamard (1865-1963) is a problem where a solution exists, the solution is unique, and the solution and the data are continuously related. As the number of sensors in EEG recordings typically are around 20-256 and the number of possible source locations are in the range between 5,000 and 1,000,000 (depending on the choice of head model and the spatial resolution of interest) no unique solution exists.

Besides the ill-posed nature of the EEG inverse problem, a serious concern is numerical instability often referred to as being ill-conditioned. Even well-posed problems may be ill-conditioned when solved with finite precision or small errors in the data, and thus, affecting the solution significantly. Evidently, speaking of ill-conditioned in relation to an ill-posed problem, this may be even more serious for the source estimates. In case of the linear problem given in (2.1) we can measure how ill-conditioned the problem is. How ill-conditioned the problem is can be represented by its condition number given by

$$\kappa = \frac{\max \sigma(\mathbf{A})}{\min \sigma(\mathbf{A})} \quad (2.2)$$

with a σ being the singular values of \mathbf{A} . The higher the condition number is the more ill-conditioned the problem is. Figure 2.4 depicts the singular values of a forward model \mathbf{A} using the BEM head model shown in Fig. 2.3(b). Indeed, this forward model is also highly ill-conditioned as the condition number is in the order of 10^{15} . Thus, care must be taken in inverse methods to minimize the influence of numerical instability of the linear system. One way to minimize the highly ill-conditioned matrix \mathbf{A} effect on the source estimate is to add a small amount of noise to the matrix, which will result in a better conditioned matrix. This is known as applying jitter noise [Neal, 1997]. In fact jitter can be included in a model as a hyperparameter such that it serves as a regularization parameter which can be optimized, c.f. [Neal, 1997].

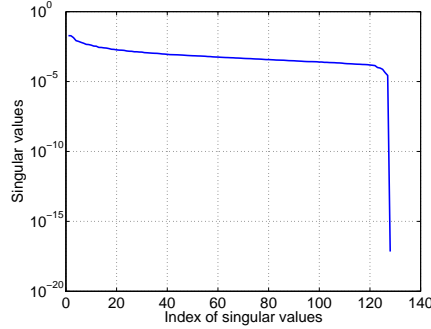


Figure 2.4: **Singular values of the lead field matrix** of a BEM head model. The condition number is in the order of 10^{15} . With the last singular value excluded a condition number of 10^3 can be obtained.

2.3 Inverse Problem

While functional brain imaging by EEG holds great promise for the neuroimaging community with its high temporal resolution and its capability of being performed under relative realistic conditions, EEG based brain imaging suffers from a lack of spatial specificity. This lack of spatial specificity is caused by the complex propagation of neural quasi-static electric fields to the array of sensors placed at the scalp surface. Motivated by the desire to perform reliable and precise reconstruction of the neural current density, much effort has been devoted to development of improved inversion methods. General approaches to the EEG inverse problem may be categorized as parametric or imaging methods [Baillet et al., 2001]. In a parametric setting, or sometimes referred to as scanning methods, the EEG is described by a small number of dipoles [Scherg et al., 1999, Moshier et al., 1992]. The source locations are found by inspecting the whole brain in order to find the best set of sources to represent the data. Choosing the number, locations, and orientations of the dipoles is a hard combinatorial problem. In contrast, imaging methods also known as distributed models reconstruct a spatial distribution of the current sources. To obtain reasonable spatial resolution, the number of source locations will necessarily be much larger than the number of sensors ($N_d \gg N_c$). Consequently, the inverse problem, estimating \mathbf{S} given \mathbf{M} is severely underdetermined (ill-posed), and all efforts at source reconstruction are heavily dependent on the regularization of the sources, i.e. which prior assumptions are made. This may be viewed as either explicitly or implicitly, as placing some prior distribution on \mathbf{S} . Additionally, existing inverse algorithms face the challenge of the many sources of noise that interfere with the true signals in the M/EEG data. Electrical, thermal, and biological noise as well as background room interference can be present. As

a consequence of the many noisy contributions and the highly ill-posed nature of the EEG inverse problem this leads to high requirements on robust inverse methods. To accomodate this need many regularization schemes have been advocated, differing by the specific prior information, typically smoothness or sparsity. The prior information can be formulated as statistical priors or mathematical constraints like the (weighted) minimum norm [Hämäläinen and Ilmoniemi, 1994] and maximum smoothness method [Pascual-Marqui et al., 1994] which assumes the sources to be diffuse and highly distributed due to their use of the l^2 -norm [Fuchs et al., 1999, Uutela et al., 1999]. In contrast models based on l^1 -norm, l^p -norms [Auranen et al., 2005], $l^{1,2}$ -norms [Haufe et al., 2011], minimum variance beamformer [Veen et al., 1997], multiple priors models, variable resolution electromagnetic tomography (VARETA) [Valdés-Sosa et al., 2000] can implement more focal sources. Other priors based on anatomy, physiology, and temporal information [Baillet and Garnero, 1997, Phillips et al., 2002, Sato et al., 2004, Trujillo-Barreto et al., 2008] have also been proposed to increase the precision of the source estimates.

One thing that all the inverse methods have in common, is that they face the challenge of determining how much regularization is needed. I.e. when the inverse models incorporate a regularization parameter into their mathematical models this parameter needs to be optimized somehow. In general the choice of regularization parameter will have a significant impact on the source solution as a too big regularization parameter will force the source solutions towards zero whereas no regularization will leave the source solution prone to errors. Often used techniques for selecting appropriate regularization parameters are, heuristic approaches such as cross-validation, the L-curve approach and the discrepancy principle [Hansen, 1992, Hansen, 2010]. As an alternative, a Bayesian formulation can be applied and the regularization parameters can be optimized jointly with the source estimate. We focus the rest of this thesis on inverse models formulated from the principle of Bayes theorem.

2.3.1 Bayesian Inference

Embedding prior knowledge into source localization models is conveniently carried out with a Bayesian framework in terms of prior probabilities of the current sources. In a Bayesian framework, probability models are used to describe the observed processes, by making inferences from data, [Gelman et al., 2003]. If a set of models are considered and we seek the most likely model that describe the data best, we can make use of Bayes' theorem. Using Bayes' theorem the posterior distribution over the model \mathcal{M} and given the observed data set \mathcal{D} the

posterior distribution spells

$$p(\mathcal{M}|\mathcal{D}) = \frac{p(\mathcal{D}|\mathcal{M})p(\mathcal{M})}{p(\mathcal{D})}, \quad (2.3)$$

where $p(\mathcal{M})$ is the prior probability distribution, whose purpose is to include a description of the uncertainty about the models that may exist. The so-called *model evidence* $p(\mathcal{D}|\mathcal{M})$ is an interesting term in Eq. (2.3). It can be viewed as a *likelihood function* over the possible models, in which the parameters have been marginalized out. Thus, it is also known as the *marginal likelihood*. The model evidence term expresses the preference of the data given a model. The denominator describes the distribution over the dataset and ensures that the left hand side results in unity when integrated over the model \mathcal{M} . Model selection is performed by selecting the model that maximizes the posterior distribution in Eq. (2.3). In contrast to model selection, where only the most probable model is used, model averaging can be performed. In model averaging, a weighting of the individual posterior probability distributions for the models is performed. Following a Bayesian perspective we are in principle advised to perform model averaging when we seek the most probable description of the underlying processes in the observed data set \mathcal{D} as this in average should lead to the best performance, [Trujillo-Barreto et al., 2004]. However, depending on the specific application model averaging may not be desirable, while averaging over a number of models can make the final model hard to interpret as each model will contribute differently to the final model.

In many cases one model is not preferable over another, and thus we apply a flat prior distribution $p(\mathcal{M})$. In this situation, maximizing the posterior distribution can be cut down to maximizing the marginal likelihood if we just seek the most probable model. For a given model \mathcal{M} the posterior distribution can be computed over the parameters in that model, by using Bayes' theorem once more

$$p(\boldsymbol{\theta}|\mathcal{D}, \mathcal{M}) = \frac{p(\mathcal{D}|\boldsymbol{\theta}, \mathcal{M})p(\boldsymbol{\theta}|\mathcal{M})}{p(\mathcal{D}|\mathcal{M})} \quad (2.4)$$

where $\boldsymbol{\theta}$ denotes the unknown parameters. The numerator consists of the likelihood function, at times merely referred to as the likelihood of the parameters, $p(\mathcal{D}|\boldsymbol{\theta}, \mathcal{M})$, and the parameter prior $p(\boldsymbol{\theta}|\mathcal{M})$ for the model \mathcal{M} . Furthermore, the denominator is represented by the marginal likelihood, whose purpose in this equation is to ensure that the left hand side results in unity when integrated over the total parameters space.

$$p(\mathcal{D}|\mathcal{M}) = \int p(\mathcal{D}|\boldsymbol{\theta}, \mathcal{M})p(\boldsymbol{\theta}|\mathcal{M})d\boldsymbol{\theta} \quad (2.5)$$

Indeed, it is seen from the above description that the marginal likelihood plays a crucial role in Bayesian inference, due to its multiple functionalities, i.e. a model selection term in Eq. (2.3) and a normalization term in Eq. (2.4). By integrating out the parameters, this penalizes models with more degrees of freedom. However, a trade-off between too complex models and too simple ones exists, since the simple ones can a priori only model a small range of data sets compared to complex ones. Too simple models may therefore not be able to capture the essential structure in the data.

Unfortunately, the marginal likelihood is often analytically intractable for most models of interest, and thus approximation is needed. The intractability arises from the fact that the integral over the parameters can be very complicated and high dimensional with couplings between parameters. One possible solution to this problem is the Variational Bayesian (VB) method, which optimizes a lower bound of the marginal likelihood. However, a number of other approximate approaches to this problem exist as well, e.g. Bayesian Information Criteria (BIC), the Laplace approximation, Expectation Propagation (EP), the Cheeseman-Stutz (CS) approximation, and Markov Chain Monte Carlo (MCMC) sampling can be mentioned, [Minka, 2001, Beal, 2003, Bishop, 2006, Oppen et al., 2008]. A review of methods for approximating the marginal likelihood can be found in [Beal, 2003], wherein it is demonstrated that VB has superior performance over BIC and CS while simultaneously keeping the computational burden relatively low.

In this thesis a VB approach will be applied. The concept of VB is to approximate the posterior distribution and the model evidence by a change of the marginal likelihood maximization problem into a maximization of the free energy over the distributions that approximates the true posterior distribution. Hereby, a lower bound is obtained on the marginal likelihood that should be maximized or equivalent minimizing the Kullback-Leibler (KL) divergence between variational and true posterior distributions, since the marginal likelihood can be decomposed into these two terms. In fact EP is quite related to VB, as EP also minimizes the KL divergence, however, just in a reversed form. EP minimizes the KL divergence of the true posterior distribution and the variational. Despite their similarities they lead to quite different properties, e.g. see [Bishop, 2006] for a further description.

Variational Bayes

Assume for a moment that the parameter set θ consists of the variables \mathbf{X} and \mathbf{Z} for describing an observed data set $\mathcal{D} = \mathbf{M}$, such that the likelihood is expressed as $p(\mathbf{M}|\theta)$. Additionally, assume that the marginal posterior distributions for the variables \mathbf{X} and \mathbf{Z} are analytically intractable. Thus, approximations are needed. Here, we apply a standard VB framework [Attias, 2000, Bishop, 2006, Ghahramani and Beal, 2000], in which a parameterized simpler distribution

$q(\boldsymbol{\theta})$ approximates the true joint posterior $p(\boldsymbol{\theta}|\mathbf{M})$. In the VB framework the parameters of the $q(\boldsymbol{\theta})$ distribution are determined with a maximization of a lower bound of the marginal likelihood expressed in terms of the Jensen's inequality,

$$\begin{aligned}\mathcal{L} &= \ln \int p(\mathbf{M}|\boldsymbol{\theta}) p(\boldsymbol{\theta}) d\boldsymbol{\theta} \\ &\geq \int q(\boldsymbol{\theta}) \ln \frac{p(\mathbf{M}|\boldsymbol{\theta}) p(\boldsymbol{\theta})}{q(\boldsymbol{\theta})} d\boldsymbol{\theta} = \mathcal{F}_{\mathcal{M}}(q(\boldsymbol{\theta})).\end{aligned}\quad (2.6)$$

$\mathcal{F}_{\mathcal{M}}$ is the lower bound for model \mathcal{M} and $\boldsymbol{\theta}$ is the union of the parameter sets $\boldsymbol{\theta} = \{\mathbf{X}, \mathbf{Z}\}$ such that the joint distribution $p(\mathbf{M}, \boldsymbol{\theta})$ is given by

$$p(\mathbf{M}, \boldsymbol{\theta}) = p(\mathbf{M}|\mathbf{X}, \mathbf{Z}) p(\mathbf{X}) p(\mathbf{Z}) \quad (2.7)$$

In order to maximize the lower bound of the log marginal likelihood we assume factorization in the parameters $\boldsymbol{\theta} = \{\mathbf{X}, \mathbf{Z}\}$ of the variational posterior $q(\boldsymbol{\theta})$, i.e., the full variational posterior distribution $q(\boldsymbol{\theta})$ is restricted to

$$q(\boldsymbol{\theta}) = q(\mathbf{X}) q(\mathbf{Z}). \quad (2.8)$$

The factorized variational distribution may be a crude approximation for the given model of interest. It is mainly motivated by computational convenience as such approximation provides tractable marginal posterior distributions. In turn each of the variational marginal posterior distributions are maximized by solving $\delta\mathcal{F}_{\mathcal{M}}/\delta q(\mathbf{X}) = 0$ and $\delta\mathcal{F}_{\mathcal{M}}/\delta q(\mathbf{Z}) = 0$, respectively. The variational posterior distributions for e.g. \mathbf{X} can then be obtained through [Attias, 2000],

$$q(\mathbf{X}) \propto e^{\langle \ln p(\mathbf{M}, \boldsymbol{\theta}) \rangle_{q(\boldsymbol{\theta}_{\setminus \mathbf{X}})}}, \quad (2.9)$$

where $q(\mathbf{X})$ denotes one of the variational marginal posterior distributions in Eq. (2.8) and the expectation $\langle \cdot \rangle$ is taken w.r.t. $q(\boldsymbol{\theta}_{\setminus \mathbf{X}})$ with $\boldsymbol{\theta}_{\setminus \mathbf{X}}$ denoting the set $\boldsymbol{\theta}$ except \mathbf{X} . Similiar the variational posterior distribution for \mathbf{Z} can be obtained.

Source Confusion

While the existing literature concerns inversion under the assumption that the forward propagation model is known, we are interested in examining this assumption in more detail in this and the following chapter. We first analyze the role of the forward model in relation to the inverse problem. More specifically we are interested in questions such as, how sources located in different brain areas are confused with other sources? And how is this affected by an uncertain forward model? In a broader perspective this analysis may be used as guidance for paradigms where EEG can be used as a reliable functional imaging tool. Furthermore, the analysis motivates us to include the uncertainty of the head model into a source reconstruction method and, thus, it can be regarded as a motivation for the inverse method **SO**urce and **FO**ward **MO**del **RE**construction (**SOFOMORE**) presented in Chapter 4. The results presented in this chapter are extensions of the contribution given in Abstract E.

3.1 Qualitative Analysis of the Role of Forward Model Uncertainty

Several studies of the influence of the tissue conductivity have already been performed [[Gencer and Acar, 2004](#), [Haueisen et al., 1997](#), [Wolters et al., 2006](#)],

with conductivity ratios between brain and skull in the range from 15 [Oostendorp et al., 2000] to 80 [Homma et al., 1995, Rush and Driscoll, 1969]. The analysis performed in this section includes BEM and spherical head models, see Fig. 2.3(a) and 2.3(b).

In order to reveal the influence of the forward model on the reconstruction of the sources, we will evaluate how sources located in different regions of the brain are confused. We analyze a single time point, thus the linear model reads $\mathbf{m} = \mathbf{A}^* \mathbf{s}^* + \boldsymbol{\varepsilon}$, where $'^*'$ indicates the ‘true’ forward model and source distribution, respectively. We will assume that the true solution is sparse, i.e. only one source, say the i^{th} , is active with a strength s_i^* . When we infer the source signals we consider the quadratic cost function with a potentially incorrect forward model $\mathbf{A} \neq \mathbf{A}^*$,

$$E(\mathbf{s}) = \|\mathbf{m} - \mathbf{A}\mathbf{s}\|_2^2. \quad (3.1)$$

The sparse least squares (sls) estimate of a single dipole solution located at the site j is $s_j^{\text{sls}} = \mathbf{a}_j^T \mathbf{m} / |\mathbf{a}_j|^2$ and zero for all other components. Inserting this into the cost function and averaging over the noise distribution $\sigma_\varepsilon^2 = \langle \boldsymbol{\varepsilon}^T \boldsymbol{\varepsilon} \rangle / N_c$ we get the expected cost:

$$\begin{aligned} \langle E(j|i) \rangle &= \left\langle \|\mathbf{m} - \mathbf{A}\mathbf{s}^{\text{sls}}\|_2^2 \right\rangle = |\mathbf{a}_i^* s_i^*|^2 \sin^2 v_{j,i} + N_c \sigma_\varepsilon^2 \\ &= |\mathbf{a}_i^* s_i^*|^2 \left(\sin^2 v_{j,i} + \frac{1}{\text{SNR}} \right) \end{aligned} \quad (3.2)$$

with

$$\cos v_{j,i} = \frac{\mathbf{a}_j^T \mathbf{a}_i^*}{|\mathbf{a}_j| |\mathbf{a}_i^*|}$$

and the signal-to-noise ratio (or inverse effective noise level) is defined as $\text{SNR} = \frac{|\mathbf{a}_i^* s_i^*|^2}{N_c \sigma_\varepsilon^2}$. Thus, even if we use the correct forward model $\mathbf{A} = \mathbf{A}^*$, hence $\sin v_{i,i} = 0$, the small signals from sulci and from cortical regions at larger distance from the sensors are more likely to be confused because the differences in angular factors $(\sin v_{j,i})^2$ can be small compared to the effective noise level. The geometrical nature of the confusion depends on the distribution of $\sin^2 v_{j,i}$ across the cortical surface. In the following we will inspect this distribution and the patterns of confusion it induces with and without uncertainty of the forward model.

The results in the analysis depend on both the cortical resolution and the electrode configuration. The dimension of \mathbf{A} and \mathbf{A}^* was set to (128×8196) . The forward propagation is based on a BEM model with conductivities $\rho_{\text{brain}} = 0.33\text{S/m}$, $\rho_{\text{skull}} = 0.0041\text{S/m}$, and $\rho_{\text{scalp}} = 0.33\text{S/m}$ corresponding to the ratios 1:1/80:1 as in [Homma et al., 1995].

In Fig. 3.1 we first analyze the confusion in case of a correct forward model. Figure 3.1(b) illustrates the confusion of the different dipoles given the sources

located in Fig. 3.1(a). For the true dipole located in the left frontal lobe (Source 1) it is seen that dipoles close to the true dipole have low costs, hence, are likely to be selected in face of noise. This is a favorable situation in which a true source in this location will be confused mainly with nearby locations. However, if the true source is located in the temporal lobe (Source 2), the confusion is amplified and distributed across large portions of cortex as seen in Fig. 3.1(d). In the following we inspect the influence of using the ratios 1:1/80:1 (the ‘correct’ ratios in the simulations) with the ratios 1:1/15:1. There are many ways to summarize the distribution of confusion. Here we consider a source properly located if the identified source is within a certain range d of the true. We can quantify the decision process using the positive predictive value (PPV) (or precision [Fawcett, 2006]):

$$\text{PPV} = \frac{\text{TP}}{\text{TP} + \text{FP}}, \quad (3.3)$$

where we define the true positives (TP) as the locations where $(\sin v_{j,i})^2 \leq 1/\text{SNR}$ and the distance to the true dipole d_{ij} is smaller than the threshold (d). A false positive (FP) is declared if $d_{ij} \geq d$ and with $(\sin v_{j,i})^2 \leq 1/\text{SNR}$. We map the distribution of PPV, and we will indicate by a white color the special cases when no occurrence of neither TP nor FP is found. In Fig. 3.1(b) the noise variance is $\sigma_\epsilon^2 = 220$, which corresponds to a signal-to-noise ratio $\text{SNR} = 10$ for a typical source (mean of all $|a_i^*|^2$). Note that depending on the location of the true source we will have different effective SNRs. It is seen that with the given noise level a source located in the temporal lobe (Source 2) is highly confused. In Fig. 3.2 the ‘true’ conductivities have been used. Figures 3.2(a) and 3.2(b) show the PPV distributions, while in Figs. 3.2(c) and 3.2(d) show worst case scenarios in two different views. In conclusion, the confusion of the reconstruction is very dependent on the location of the true source. It is well-known that sources located in sulci are generally more difficult to reconstruct in EEG compared to sources at gyri, since the orientation of the sources from most of the sulcus areas is tangential. As we will see in the following these problems are further amplified when an incorrect forward model is applied.

Forward fields with conductivity errors

A set of brain images for the case with brain:skull conductivity ratio 1:1/15 can be found in Fig. 3.3, providing some insight in the role of conductivity uncertainty, c.f., Fig. 3.2. Indeed, a wrong conductivity ratio influences our ability to reconstruct sources correctly. Note the increase in areas where neither TP nor FP (occipital and parietal regions) have been found, which indicates an increase in the angular factors. In contrast to the case of reconstruction with a true brain:skull conductivity ratio (1:1/80) some areas like motor cortex and the lower part of the brain actually seems to be less confused when using the erroneous conductivity ratio 1:1/15! However, the angular factors also increase indicating a poorer representation of the signal. That the angular factors have

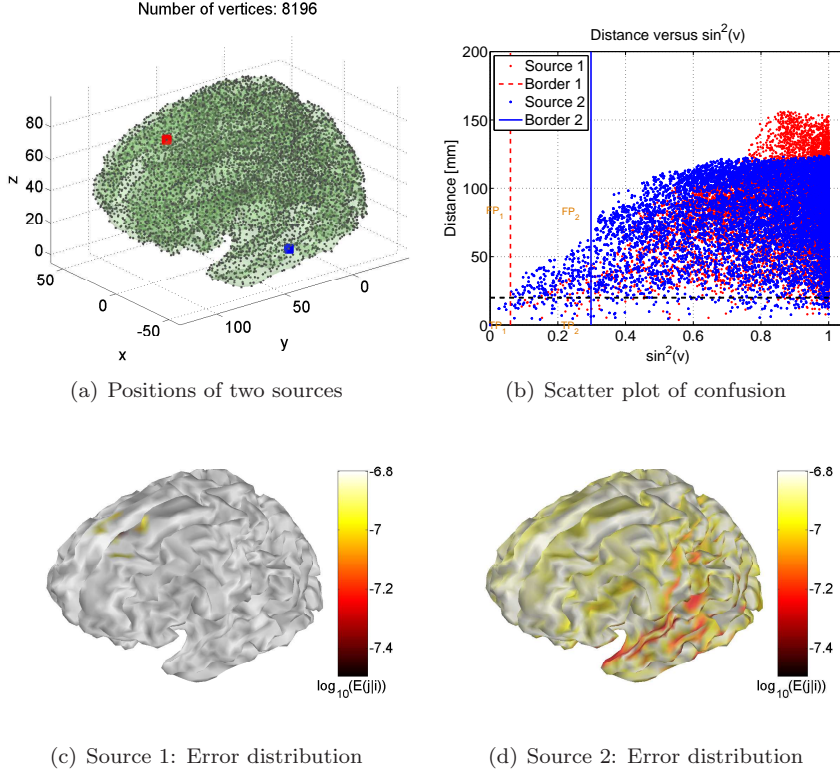


Figure 3.1: Comparison of the amount of **expected confusion for two sources** located in different cortical regions. (a) The positions of the two sources that are compared. Source 1 (left frontal part) and Source 2 (left temporal lobe). (b) Scatter plots of the distance d_{ij} between the true source and each of the candidates (j^{th} dipole) versus angular factors $(\sin v_{j,i})^2$. The colors of the dots correspond to their true source Source 1 & 2 in the top left plot. On the left side of Border 1 & 2 we have $(\sin v_{j,i})^2 \leq 1/\text{SNR}_1$ and $(\sin v_{j,i})^2 \leq 1/\text{SNR}_2$, respectively, where SNR_1 denotes the SNR when Source 1 is the true source and similar if we have Source 2. Likewise, subscript 1 and 2 for TP and FP are also related to the true sources Source 1 and Source 2. The threshold that separates TP and FP is $d = 20\text{mm}$. (c)+(d) The log error distribution illustrated on the cortex with Source 1 and Source 2 as the true source, respectively.

increased is supported by Fig. 3.4.

Forward fields with geometric errors

In Fig. 3.5 a forward model with ‘errors in the geometry’ is analyzed. Geometric

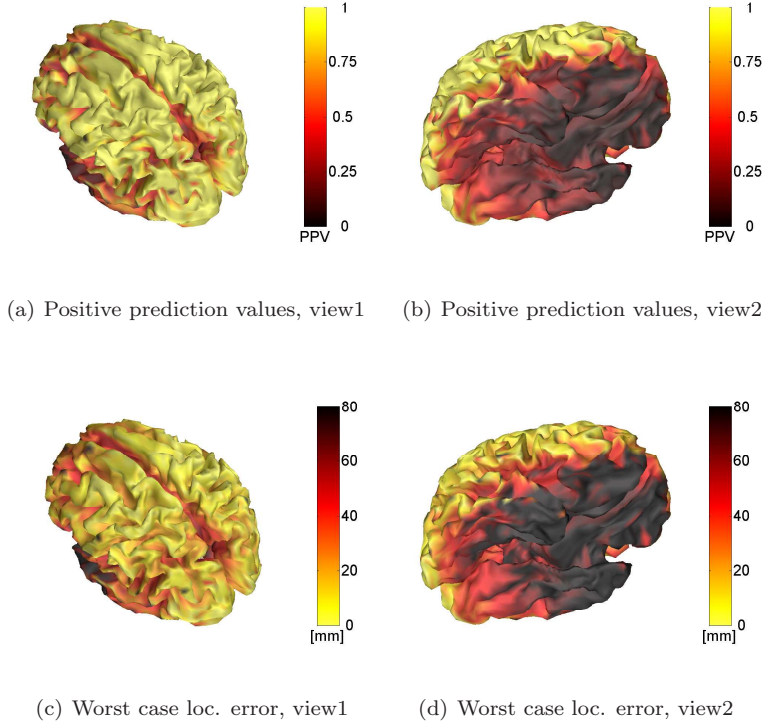


Figure 3.2: Confusion measured in terms of **PPVs** and **localization errors** for the different parts of the cortex **given the ‘correct’ conductivity ratios** $\rho_{\text{brain}} : \rho_{\text{skull}} : \rho_{\text{scalp}} = 1 : 1/80 : 1$, $d = 20\text{mm}$, and mean $SNR = 10$. The occipital and the frontal lobe of the brain have quite high precision at the given noise level. In contrast the lower parts of the brain, like the temporal lobe, have much lower precision at the given noise level. Sources located in sulci are in general more confused than gyral sources.

errors are induced by applying a 3-spheres head model in stead of a BEM as in Fig. 3.2. The consequence of such geometric errors in \mathbf{A} is a general decrease in the PPV’s and an increase in localization error especially for the upper part of the brain, e.g., the parietal lobe. Similar to the inspection of minimum angular factors when conductivity errors are present an analysis of the minimum angular factors when geometric errors are present reveals an increase in the angular factors, c.f. Fig. 3.6. With the angular factors increasing this means a poorer representation of the signal.

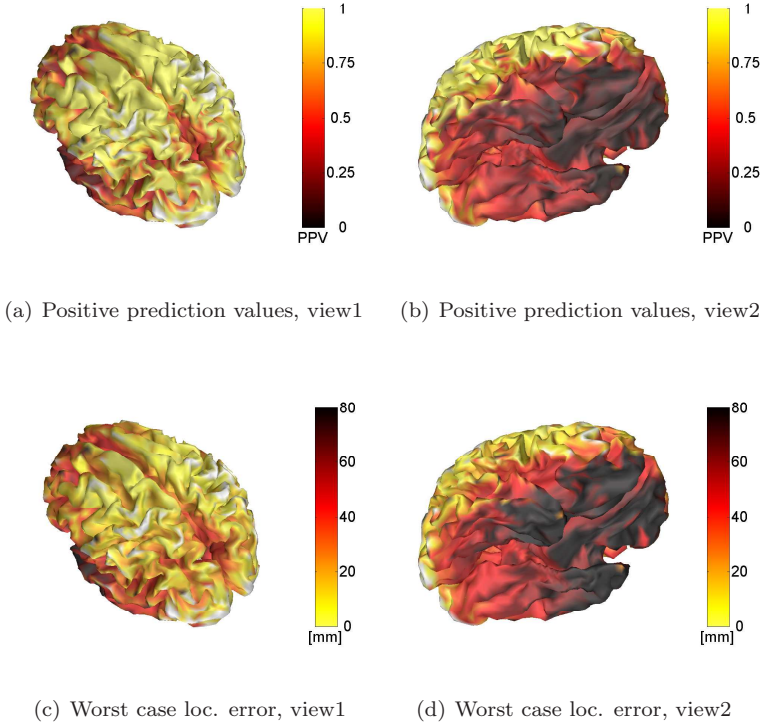


Figure 3.3: **PPVs and localization errors** for the different parts of the cortex given the conductivity ratios $\rho_{\text{brain}} : \rho_{\text{skull}} : \rho_{\text{scalp}} = 1 : 1/15 : 1$, $d = 20\text{mm}$, and mean $SNR = 10$. Increase in white areas where no TP or FP have been found (occipital and parietal regions) indicating increase in angular factors.

3.2 Related Work

Analyzing the importance of precise forward models and their influence on either the forward problem or the inverse problem is far from new to the EEG community. A number of contributions have already been published e.g. [Huiskamp et al., 1999, Vanrumste et al., 2000, Oostenveld and Oostendorp, 2002, Gencer and Acar, 2004, Ramon et al., 2006, Wolters et al., 2006]. A majority of the forward model investigations performed evaluates the forward models at sensor level or only examine a few dipoles located in different brain regions. These contributions have been of crucial importance to the EEG community as they have shed light on serious issues that we need to be aware of when the source solutions are used as the basis of conclusions in a given setup. In [Huiskamp et al., 1999]

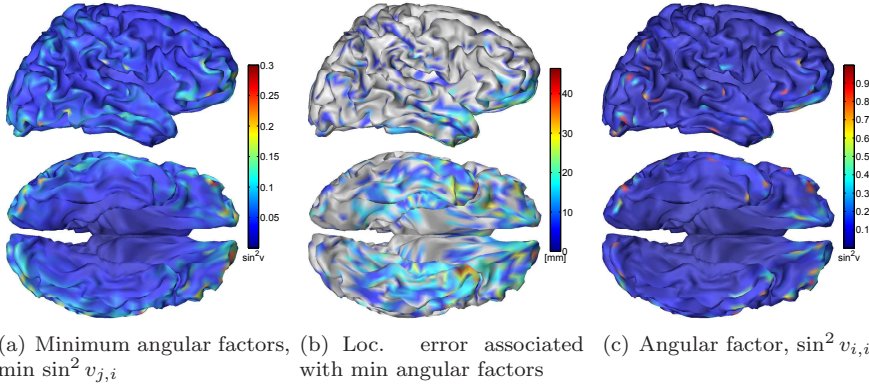


Figure 3.4: Increase in **angular factors** as a consequence of **mismatch in the skull conductivity value**.

an examination of the influence of geometric errors on the source estimates is performed using BEM models constructed from MR or CT images. In contrast to [Huiskamp et al., 1999] we do not restrict the analysis to a few dipoles but instead explore how each dipole out of a set of 8196 dipoles distributed over the cortex are confused with each other. A quite similar approach as the analysis above have been presented in [Fuchs et al., 1998]. The study in [Fuchs et al., 1998] deals with geometric errors introduced by using too simple head models (spheres) compared to more realistic BEM models. To evaluate the effect, angles between forward fields were examined. However, in contrast to [Fuchs et al., 1998], the analysis given here and in Abstract E explores forward model uncertainties in face of noise present at the sensor level. From Eq. (3.2) it is clear that the source confusion is indeed dependent on the interplay of forward field errors and the amount of noise present in the recordings. Other types of geometric errors such as not taking holes in the skull into consideration have been examined by [Vanrumste et al., 2000, Oostenveld and Oostendorp, 2002]. In [Vanrumste et al., 2000, Oostenveld and Oostendorp, 2002] it is demonstrated that if holes are excluded from the head model this will lead to a significant decrease in the ability to reconstruct the sources correctly.

Errors associated with the tissue conductivity values and the importance of modeling specific tissues as anisotropic rather than isotropic have been discussed in [Gencer and Acar, 2004, Ramon et al., 2006, Wolters et al., 2006], with the overall consensus that inaccurate modeling of the skull leads to significant error contributions on the sources. In fact [Wolters et al., 2006, Hallez et al., 2007] states that the a smearing effect on the forward potential computation is introduced by the skull anisotropy. The deeper a source is located the more it is surrounded by anisotropic tissues. Thus, electric fields generated by deeper

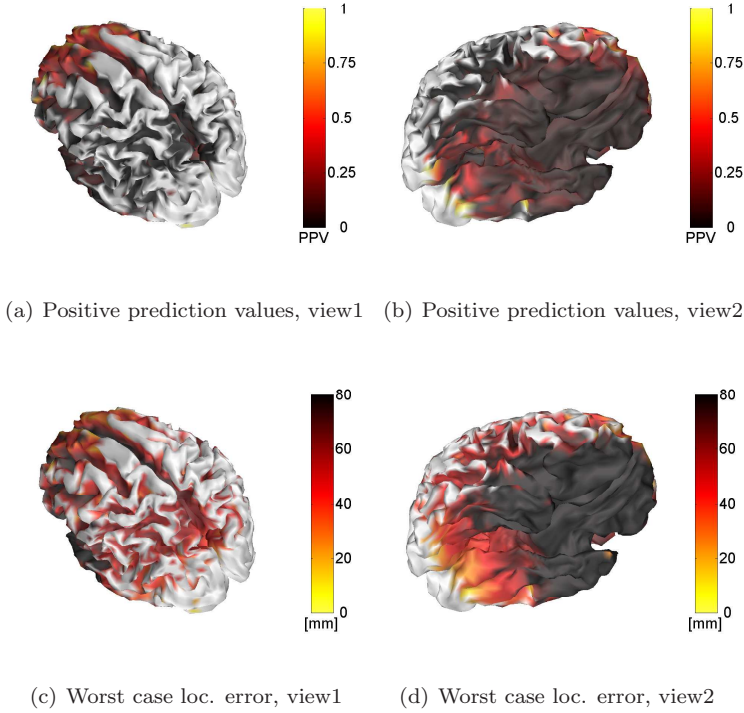


Figure 3.5: **PPVs and localization errors** for the different parts of the cortex given the conductivity ratios $\rho_{\text{brain}} : \rho_{\text{skull}} : \rho_{\text{scalp}} = 1 : 1/80 : 1$, $d = 20\text{mm}$, mean $SNR = 10$, and **geometric errors** due to the usage of a 3-spheres model.

sources are more affected by the anisotropy than superficial sources.

Of more recent studies [Chen et al., 2010, Steinsträter et al., 2010] should be mentioned, which apply a quite similar approach as given in the analysis above. Especially, [Steinsträter et al., 2010] performs a careful analysis of how the number of electrodes, geometric errors (spheres versus FEM head models), and anisotropy versus isotropy affect the source estimates obtained by beamformers.

3.3 Summary

The source confusion analysis performed here revealed how errors in forward model may deteriorate the source estimates as sources are confused with each

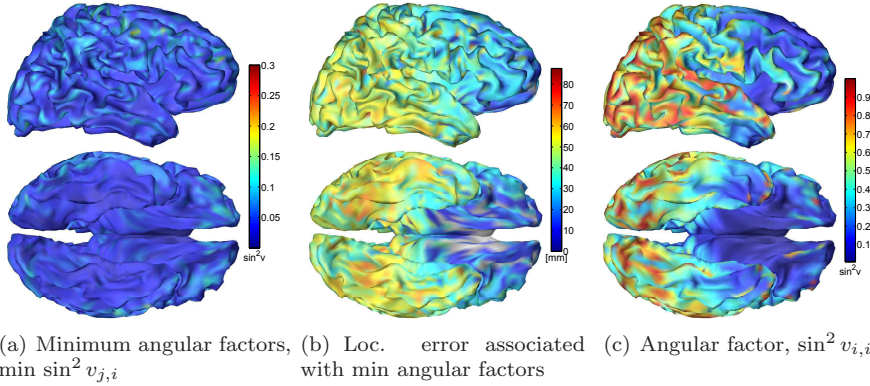


Figure 3.6: Increase in angular factors as a consequence in geometric differences between BEM and 3-spheres head models.

other when noise is present in the EEG recordings. Thus, in order to minimize the risk of sources to be confused with each other an accurate forward model is needed. However, it should be noted that even with the true forward fields the sources are likely to be confused with each other in face of noise due to the fact that we are dealing with an inverse problem. The degradation of the source estimates were demonstrated from an analysis of simple sparse least squares solutions by evaluating angular factors between 'true' forward fields and potential 'erroneous' forward fields. As no spatial or temporal regularization accompanied this simple inverse method it must be expected that more advanced inverse methods will minimize source confusion. Even though this analysis is based upon a simple inverse method it may serve as a worst case scenario for what can be expected when errors in the forward model is present.

Besides the choice of head model another type of geometric error that can appear in forward models is if no individual information about a subjects head, such as if no structural MRI or CT scan are available. From a clinical perspective it is far from everyone that has a MRI or CT scan of their brain and thus it is problematic to construct realistic individual head models for such subjects using BEM, FEM or FDM models as they will be deemed quite uncertain in terms of a mismatch between a template head and a subject specific head. To the authors knowledge no current careful evaluation of the influence on the source estimates exists when using a template head model rather than a realistic head model. However, it should be straight forward to do as well, since the approach taken in this chapter or alternatively as in [Steinsträter et al., 2010] can be applied.

Source Localization in Forward Model Uncertainty

In light of the serious source confusion which may exist due to errors that are present in the forward model, an alternative/addition to the advanced head models is to minimize the effect of source confusion when performing the source estimation. Therefore, we seek a joint optimization of a forward model correction and source estimation. This section describes the proposed method - **S**ource and **F**orward **M**odel **R**Econstruction - in short **SOFOMORE** (see contributions **A**, **B**, and **C**). **SOFOMORE** treats the forward model as a stochastic process. In contrast, existing source localization methods normally assume the forward model as known and fixed. Modeling the forward model as being stochastic is motivated by the many noise processes that contribute to the forward model, including the geometrical representation of the cortical surface, the conductivity distribution, and electrode positions. When 'realistic head models' are constructed from tissue segmentation based on e.g. structural MRI, the geometry of the head model is affected by the resolution and tissue segmentation errors. These errors can be presented as small magnitude perturbations of the head model shape [von Ellenrieder et al., 2006].

An illustration of the concept of applying a stochastic forward model in **SOFOMORE** is given in Fig. 4.1, in which the conventional forward model $\mathbf{A}^{(0)}$ is represented as 'the prior mean'. Due to the uncertainties mentioned above,

we expect a difference between the true underlying model of the forward fields, \mathbf{A}^* and $\mathbf{A}^{(0)}$. The main novelty of our SOFOMORE approach is thus the posterior distribution of the forward model \mathbf{A} representing our uncertainty of the deviations of the prior mean $\mathbf{A}^{(0)}$.

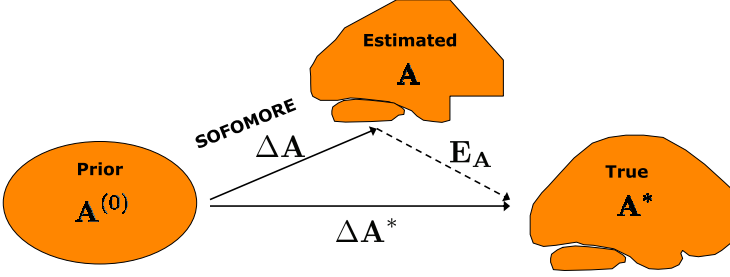


Figure 4.1: Illustration of **the concept of forward model reconstruction**. The SOFOMORE model uses a potential erroneous forward model $\mathbf{A}^{(0)}$ as prior mean and depending on the regions of source activity, the model can flexibly correct these regions by $\Delta\mathbf{A}$. As such, the estimated forward model is \mathbf{A} with error \mathbf{E}_A to the true underlying forward model denoted \mathbf{A}^* .

We here focus on the conceptual development and the statistical aspects of the SOFOMORE model.

4.1 Method: The SOFOMORE Model

Given the linear relation between the sensors and the current sources within the brain in Eq. (2.1) the observation model can be expressed as, $p(\mathbf{m}_t | \mathbf{s}_t, \Sigma_{\mathcal{E}}) = \mathcal{N}(\mathbf{m}_t | \mathbf{A}\mathbf{s}_t, \Sigma_{\mathcal{E}})$ where the noise is assumed to be time independent Gaussian distributed with $\Sigma_{\mathcal{E}}$ representing the noise spatial covariance matrix. If needed temporal correlations can be incorporated as outlined in [Friston et al., 2008]. In the remainder of this chapter the noise covariance matrix is further simplified with $\Sigma_{\mathcal{E}} = \beta^{-1}\mathbf{I}_{N_c}$ and will be estimated in the model. Alternatively, a number of approaches e.g. [Nagarajan et al., 2007, Wipf et al., 2010], estimate the noise covariance matrix from pre-stimulus data when this is available in event-related experiments. Estimating a full noise covariance matrix can be problematic depending on the amount of data available. In general this is a concern for all models comprising spaces of high dimensionality. In fact they underly what is known as the curse of dimensionality, meaning increasing the dimensionality

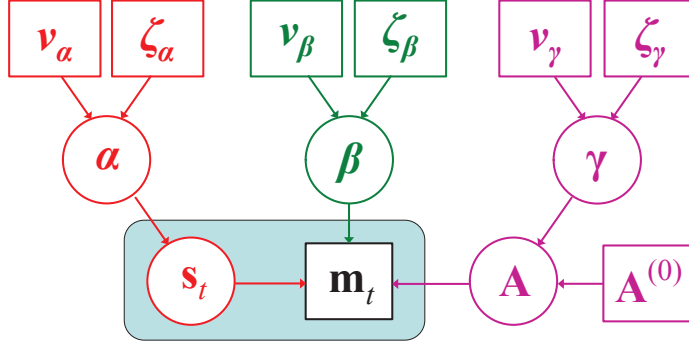


Figure 4.2: **Graphical representation of the SOFOMORE model.** The blue box including the sources \mathbf{s}_t and observations \mathbf{m}_t indicates expansion over time t . At the lowest level in the hierarchical structure we also find the forward model \mathbf{A} with fixed prior mean $\mathbf{A}^{(0)}$. The middle layer includes α precision parameter for the sources with a separate precision parameter (inverse variance) assigned to each dipole. β is the inverse variance of the noise contribution and γ includes a precision parameter to each column in \mathbf{A} . At the top level we have the hyperhyperparameters controlling the hyperparameters in the middle layer.

number often leads to an exponential growth in the required number of observations [Bishop, 2006].

As proposed in contribution A,B, and C, we here represent the SOFOMORE model, with a hierarchical structure expressed by

$$\begin{aligned}
 \mathbf{m}_t &\sim \mathcal{N}(\mathbf{m}_t | \mathbf{A}\mathbf{s}_t, \beta^{-1}\mathbf{I}_{N_c}) \\
 \mathbf{s}_t &\sim \mathcal{N}(\mathbf{s}_t | \mathbf{0}, \mathbf{D}^{-1}), \quad \mathbf{D} = \text{diag}(\alpha) \\
 \mathbf{a}_i &\sim \mathcal{N}(\mathbf{a}_i | \mathbf{a}_i^{(0)}, \gamma_i^{-1}\mathbf{I}_{N_c}) \\
 \alpha_i &\sim \mathcal{G}(\alpha_i | \nu_{\alpha_i}, \zeta_{\alpha_i}) \\
 \gamma_i &\sim \mathcal{G}(\gamma_i | \nu_{\gamma_i}, \zeta_{\gamma_i}) \\
 \beta &\sim \mathcal{G}(\beta | \nu_{\beta}, \zeta_{\beta}).
 \end{aligned} \tag{4.1}$$

A graphical representation of the hierarchical model is given in Fig. 4.2. At the top level we have the hyperhyperparameter $\{\nu_{\alpha}, \zeta_{\alpha}, \nu_{\beta}, \zeta_{\beta}, \nu_{\gamma}, \zeta_{\gamma}\}$ controlling the middle layer with the hyperparameters $\{\alpha, \beta, \gamma\}$, which control the lowest level of parameters $\{\mathbf{A}, \mathbf{s}_t\}$. For the current sources we assign a zero-mean multivariate Gaussian prior distribution with a diagonal precision matrix \mathbf{D} . In the diagonal we have $\alpha = \{\alpha_i\}_{i=1}^{N_d}$ such that the model allows dipoles to have

different variance. The current variances are assumed not to change over the time period $t = 1 : N_t$ and are modeled by an automatic relevance determination-type prior (ARD), using a conjugate prior for the precision parameters α . The conjugate prior for a precision parameter of a 1D Gaussian distribution is a Gamma distribution, $\mathcal{G}(\cdot)$, with shape parameter ν_{α_i} and rate parameter ζ_{α_i} such that the mean is given by $\nu_{\alpha_i}/\zeta_{\alpha_i}$ and the variance $\nu_{\alpha_i}/\zeta_{\alpha_i}^2$. Assigning an ARD prior to the sources is a key aspect in preventing overfitting, since by optimization of the hyperparameters it is possible to drive a proportion of the sources towards zero automatically. Driving parts of the sources towards zero are obtained when their precision parameters grow towards infinity indicating that these sources are deemed no relevance to the model as they do not contribute significantly to the description of the observations. See e.g. [Bishop, 2006] for a more detailed explanation of the ARD property.

As we regard the forward model as a stochastic process, the conventional forward propagation model $\mathbf{A}^{(0)}$ is used as a mean value in a multivariate Gaussian prior. Each of the forward fields are modeled as independent, such that the prior distribution of a single forward field is given by prior mean $\mathbf{a}_i^{(0)}$ (i'th column in $\mathbf{A}^{(0)}$) and precision γ_i , i.e. we have $\gamma = \{\gamma_i\}_{i=1}^{N_d}$. Assigning each of the forward fields a separate precision parameter allows us to differentiate the individual amounts of correction to the forward fields. The precision parameters for the forward fields are also modeled by conjugate prior distributions, i.e., Gamma distributions with shape and rate parameter, ν_{γ_i} and ζ_{γ_i} , respectively. Likewise, the inverse noise variance is assumed to be Gamma distributed, $\mathcal{G}(\beta|\nu_\beta, \zeta_\beta)$. The reasoning behind assigning an extra layer (the top layer) to the hyperparameters, is to obtain posterior distributions rather than point estimates. In addition, choosing a prior distribution over the precision parameters allows the model a-priori to constrain the model to a specific range of values, e.g. if we have an idea of the interval that the SNR for a given experiment should be in. Thus, in the prior distribution for β we have the possibility to specify precision values corresponding to expected SNR levels. One may argue that assuming the noise covariance matrix to be an identity matrix multiplied by a scalar is a crude approximation as the sensors will be highly correlated. Indeed this is a crude approximation if no preprocessing is carried out. However, the noise correlation between electrodes can be reduced with techniques similar to ‘whitening’ of the data, i.e. before passing the EEG data to the SOFOMORE model we can project the data into a space determined by the inverse of pre-estimated noise covariance matrix. As suggested above pre-estimation of the noise covariance in an event-related experiment can be performed by taking advantage of pre-stimulus data. Assuming such a setup the role of the β parameter is to capture any changes in the noise level that potentially could be present between pre-stimulus and post-stimulus periods.

We note that a special case of the SOFOMORE solution is the well-known minimum norm (MN) method [Hämäläinen and Ilmoniemi, 1994]. In fact SOFOMORE depicts the MN solution when the prior for the forward fields are just delta functions centered at $\mathbf{a}_i^{(0)}$ and if all the source precision values are equal.

Given the hierarchical structure of the SOFOMORE model with the parameters $\boldsymbol{\theta} = \{\mathbf{S}, \mathbf{A}, \boldsymbol{\alpha}, \boldsymbol{\gamma}, \beta\}$ the marginal posterior distribution of the current sources becomes analytically intractable, thus, approximations are needed for obtaining posterior distributions of \mathbf{S} and \mathbf{A} . Here, we apply a standard VB framework [Attias, 2000, Bishop, 2006, Ghahramani and Beal, 2000], in which a parameterized simpler distribution $q(\boldsymbol{\theta})$ approximates the true joint posterior $p(\boldsymbol{\theta}|\mathbf{M})$ as described in section 2.3.1.

In order to maximize the lower bound of the log marginal likelihood we further assume factorization in the parameters $\boldsymbol{\theta} = \{\mathbf{S}, \mathbf{A}, \boldsymbol{\alpha}, \boldsymbol{\gamma}, \beta\}$ of the variational posterior $q(\boldsymbol{\theta})$, i.e., the full variational posterior distribution $q(\boldsymbol{\theta})$ is restricted to

$$q(\boldsymbol{\theta}) = q(\mathbf{S}) q(\mathbf{A}) q(\boldsymbol{\alpha}) q(\beta) q(\boldsymbol{\gamma}). \quad (4.2)$$

The factorized variational distribution is a crude approximation and it is mainly motivated by computational convenience as it provides for a tractable marginal posterior distribution. In turn each of the variational marginal posterior distributions are maximized by evaluating the complete log likelihood given by

$$\begin{aligned} \ln p(\mathbf{M}, \boldsymbol{\theta}) = & \frac{N_t N_c}{2} \ln \beta - \frac{\beta}{2} \sum_{t=1}^{N_t} (\mathbf{m}_t - \mathbf{A} \mathbf{s}_t)^2 + \frac{N_t}{2} \ln |\mathbf{D}| \\ & - \frac{1}{2} \sum_{t=1}^{N_t} \mathbf{s}_t^T \mathbf{D} \mathbf{s}_t + \frac{N_c}{2} \ln |\mathbf{G}| - \frac{1}{2} \sum_{j=1}^{N_c} (\mathbf{l}_j - \mathbf{l}_j^{(0)})^T \mathbf{G} (\mathbf{l}_j - \mathbf{l}_j^{(0)}) \\ & + \sum_{i=1}^{N_d} \{(\nu_{\alpha_i} - 1) \ln \alpha_i - \zeta_{\alpha_i} \alpha_i\} + \sum_{i=1}^{N_d} \{(\nu_{\gamma_i} - 1) \ln \gamma_i - \zeta_{\gamma_i} \gamma_i\} \\ & + (\nu_{\beta} - 1) \ln \beta - \zeta_{\beta} \beta + \text{const}, \end{aligned} \quad (4.3)$$

where only the terms with the parameters involved are kept and the rest are assigned to the constant term ‘const’.

Following this scheme we obtain the sequential VB updates as illustrated in Fig. 4.3, in which each of the variational marginal posterior distributions of the parameters are given. A detailed derivation of the VB updates for the SOFOMORE model can be found in [Stahlhut et al., 2010b].

4.2 Empirical Evaluation

As the viability of the source estimates have already been demonstrated on both real EEG data and artificial data using a number of different validation metrics in [Stahlhut et al., 2009b, Stahlhut et al., 2009a, Stahlhut et al., 2010b] we only give a brief overview of how the SOFOMORE model performs. In order to understand the effect of regarding the forward fields as stochastic processes we compare with a similar hierarchical model, however, with a fixed forward model $\mathbf{A}^{(0)}$. I.e. this corresponds to replacing all the elements with the color purple with a fixed forward model equal the prior mean $\mathbf{A}^{(0)}$.

As illustration we report results from simulated data here as many unknown parameters are involved in real EEG data - the exact source distribution and 'correct' forward fields. In contrast to the simulations we select a forward model as the true forward model. In order to imitate that the cortex is of high resolution (actually infinite as continuous) and that we cannot resemble it perfectly, the 'true' forward model is selected to have a higher resolution than the forward models used by the inverse methods. Of artificial signal we adopt simulations setups previously used in the literature, see e.g. [Grova et al., 2006] for similar simulation setups. The artificial source signal consists of small patch with time evolution expressed by a half sine with duration 50ms from $t = 25\text{ms}$ to $t = 75\text{ms}$. The small patch that is selected as being the true source region is shown in Fig. 4.4.

The simulated clean EEG was constructed with forward model consisting of 3-spheres with tissue conductivities brain:skull:scalp = 0.33:0.0041:0.33S/m (ratio 1:1/80:1) and dimension 128×7204 . We corrupted the clean EEG with 'realistic EEG noise' obtained from the pre-stimulus periods in an evoked EEG study described in [Henson et al., 2003]. We applied a SNR = 10, with SNR defined as $\text{SNR} = P_{\text{cEEG}}/P_{\mathcal{E}}$, where P_{cEEG} and $P_{\mathcal{E}}$ are the power of the clean EEG and noise, respectively. Conductivity errors were added to one of the forward models used in the source localization.

The source estimates for the ARD and SOFOMORE method are given in Fig. 4.5. Three forward models with different conductivity values (brain:skull:scalp) are used as source:

- **A1** 0.33:0.0223:0.33S/m (ratio 1:1/15:1)
- **A2** 0.33:0.0041:0.33S/m (ratio 1:1/80:1)
- **A3** 0.33:0.0013:0.33S/m (ratio 1:1/250:1)

The cortical resolution in all three forward models is 4004 vertices. In general the ARD, reconstructs minor part of the simulated region and few other small regions. Activity of small sizes, which especially is the case for the ARD estimates is highlighted with circles around the activity to help the reader capturing these small regions. In contrast the SOFOMORE more or less capture the simulated activity. However, with reconstructed regions a bit larger than the simulated ones. Large distance errors have been reduced by the SOFOMORE compared to the ARD.

4.3 Related Work

The issue of modeling the forward process have also been pursued elsewhere [Lew et al., 2007, Plis et al., 2007], however, from quite a different perspective than the SOFOMORE approach. In [Lew et al., 2007] the geometric structure of the forward model is assumed known, while the skull-brain conductivity ratio is an unknown parameter which is estimated simultaneously with a few sources in a dipole fit setting. A follow up on the work in [Lew et al., 2007] is a probabilistic approach for distributed models proposed by [Plis et al., 2007] in which uncertainties of the skull conductivity is modeled as a variable. An alternative way to deal with forward model uncertainties has been presented in [Henson et al., 2009]. In this work a Bayesian framework is described for dealing with uncertainty related to forward models in MEG studies. Furthermore, [Henson et al., 2009] compare different types of head models single-sphere, overlapping sphere and BEM using the model evidence for each of the models. Moreover, the cortical resolution and the influence of constraining the source orientation perpendicular to the cortex were also inspected. The experiments were carried out using two different inverse methods, the MN [Hämäläinen and Ilmoniemi, 1994] and MSP [Friston et al., 2008].

In contrast to previous work in the field we have proposed a more unconstrained approach by modeling the whole forward model as uncertain using a prior distribution for the forward model. The hierarchical structure with the ARD type of priors used in the SOFOMORE model can be related to a quite similar structure used in [Sato et al., 2004]. However, [Sato et al., 2004] differs from the SOFOMORE model in the sense that they also incorporate fMRI prior information but on the other hand limit their model to a fixed forward model. As an alternative to the SOFOMORE model the total least squares (TLS) [Golub et al., 2000] can be mentioned, which is an extension of the ordinary least squares problem such that it also can deal with errors in \mathbf{A} .

4.4 Summary

We have given an overview of the concept and the hierarchical structure of the SOFOMORE model for performing simultaneous estimation of the current source density and the forward model. The proposed model may be considered as a first step for modeling uncertainty in \mathbf{A} by assuming the errors to be Gaussian distributed. The next step may be a more realistic model of the uncertainty arising from sensor placement and head shape. In such a model one would let the uncertainty in an element of \mathbf{A} depend non-linearly upon the distance from the sensor to the source [Mosher et al., 1999]. In continuation to this, a concern associated with the SOFOMORE approach is, that the changes performed to the prior mean $\mathbf{A}^{(0)}$ may have multiple interpretations as the changes performed can be related to changes in tissue conductivity values, sensor configuration, or alternatively by modifying the structure of the cortex. In fact it is difficult with the current setup to extract the physical changes that have been made to the forward model by SOFOMORE in a real EEG experiment. However, evaluation of the SOFOMORE on real EEG data in [Stahlhut et al., 2009b, Stahlhut et al., 2009a, Stahlhut et al., 2010b] has demonstrated the viability of the model as the source estimates are consistent with prior expectations in an experiment with faces as visual stimuli.

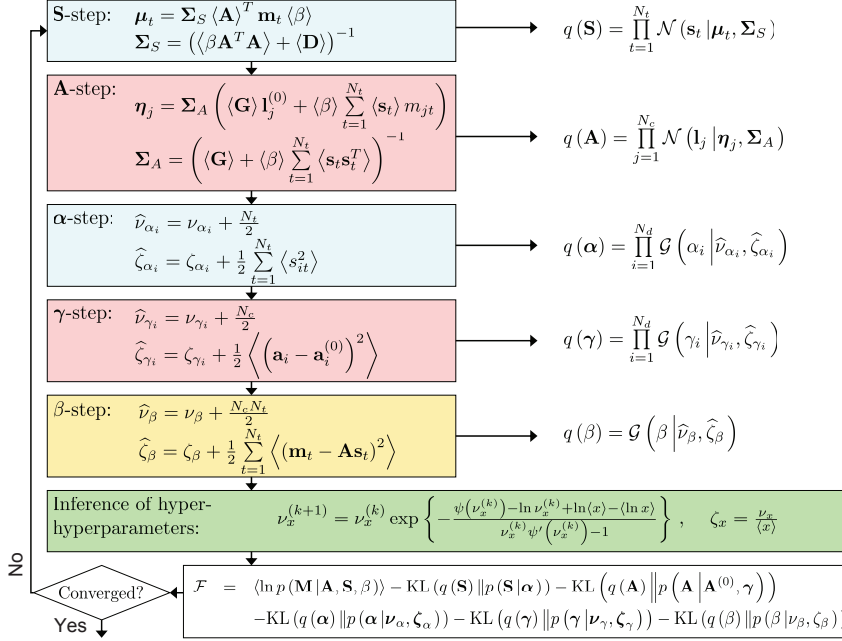


Figure 4.3: **Flow chart of the SOFOMORE model** in a VB formulation. A fully factorized model in $\boldsymbol{\theta} = \{\mathbf{S}, \mathbf{A}, \boldsymbol{\alpha}, \boldsymbol{\gamma}, \beta\}$ is applied, where each of the VB updates can be seen with their corresponding approximate posterior distribution to the right. The log evidence of the model is denoted \mathcal{F} , which include Kullback-Leibler (KL) divergences of the approximate posterior distributions and their prior distributions. Following notations are used $\langle \cdot \rangle$ denoting the expectation, \mathbf{l}_j^T the j 'th row in \mathbf{A} , and \mathbf{G} a diagonal matrix with $\boldsymbol{\lambda}$ in the diagonal. Furthermore, $\psi(\cdot)$ is the digamma function and its derivative is $\psi'(\cdot)$. Hyper-hyperparameters of the form ν_x and ζ_x are updated using Newton-Raphson. The colorcoding light blue, red, and yellow correspond to the blue, red, and yellow colorcoding to the left, i.e. the source, forward fields, and channels, respectively. The green colorcoding of the step with inference of hyper-hyperparameters means that both hyper-hyperparameters associated with sources, forward fields, and channels are updated here.

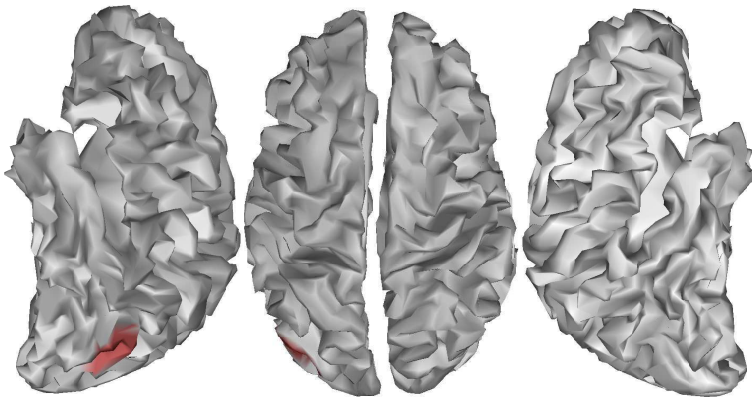


Figure 4.4: **Simulated region of source activity.** The active patch is highlighted as red.

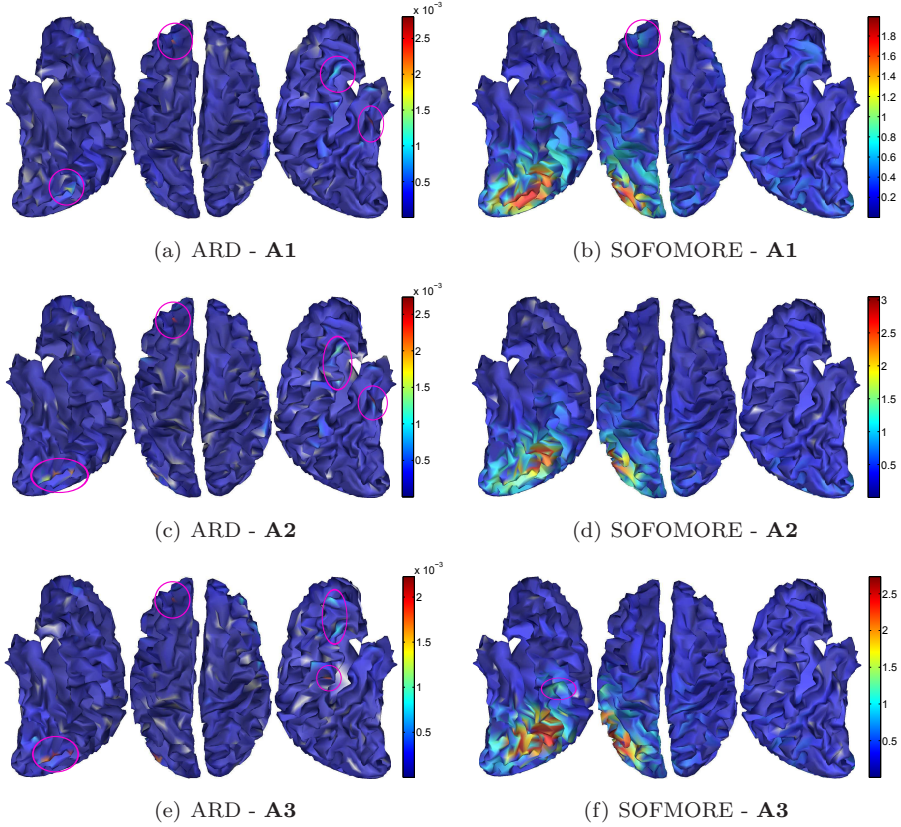


Figure 4.5: **Reconstructed source activity by ARD and SOFOMORE.** The source activity is represented by ‘Bayesian z-scores’, i.e. $s_{it}\sqrt{\alpha_i}$. Three different forward models are tested, **A1**, **A2**, and **A3**. The forward model **A2** has the same conductivity values as the true forward model **A***. In the ARD method (left column) no corrections of the forward models is performed. In the SOFOMORE method (right column) corrections of the forward models are integrated in the model. To ease the detection of small regions of activity these have been highlighted with circles.

CHAPTER 5

Spatio-Temporal Approach for Source Localization

In this chapter we discuss spatio-temporal approaches to EEG source localization and with the focus on a proposed spatio-temporal model called the Aquavit model originally presented in contribution [D](#). As we will see shortly, the Aquavit consists of temporal basis functions in order to obtain a smooth temporal representation. In this respect it is important to realize that in general a crucial aspect with spatio-temporal models is that the performance using basis functions is normally highly dependent on the choices of basis functions. There are many approaches that researchers have taken to form spatial and temporal basis functions, see e.g. [[Baillet and Garnero, 1997](#), [Daunizeau et al., 2006](#), [Efferen et al., 2000](#), [Trujillo-Barreto et al., 2008](#), [Bolstad et al., 2009](#), [Zumer et al., 2008](#), [Dyrholm et al., 2007](#)]. Fixed basis functions include Fourier series, wavelet or discrete prolate spheroidal functions which specify bandwidth and temporal extent of the data. Learned basis functions include PCA, ICA and related approaches which learn basis functions from data.

5.1 Method: The Aquavit Model

With the relation between the measured EEG signal and the brain's current sources expressed by a linear instantaneous form in the sources as in Eq. (2.1) and the full noise covariance matrix $\Sigma_{\mathcal{E}}$ we can represent the Aquavit model as the graphical model in Fig. 5.1. In this chapter we assume the noise covariance matrix as fixed and keep $\Sigma_{\mathcal{E}}$ in the formulation, even though we in principle can replace it by an identity matrix if we project it through a space spanned by the inverse noise covariance matrix before passing the recorded EEG signal to the Aquavit algorithm. The reasoning behind leaving the full covariance matrix in this formulation is to highlight its similarities to the associated algorithm called *Champagne*, [Wipf et al., 2010], which does not include temporal prior information. As previous stated the reasoning behind leaving the spatial structure of noise covariance fixed is that it is possible to suppress noisy factors (e.g., using a variational Bayesian factor model proposed in [Nagarajan et al., 2007]) effectively when pre-stimulus data is available.

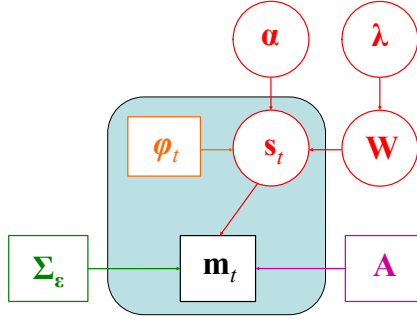


Figure 5.1: **Graphical representation of the Aquavit model.** The blue box including the sources \mathbf{s}_t , the temporal basis function φ_t and observations \mathbf{m}_t indicates expansion over time t . Besides φ_t the sources \mathbf{s}_t consist of the spatio-temporal maps \mathbf{W} and the precision parameters α . A separate precision parameter is assigned to each of the dipoles. Similar, the spatio-temporal maps are controlled by the hyperparameters λ . Finally, $\Sigma_{\mathcal{E}}$ denotes the noise covariance matrix.

From Fig. 5.1 it is noted that the current sources consist of the three components, Φ , \mathbf{W} , and α . In fact we define the current sources \mathbf{S} to encompass both temporal evolution and source noise such that the source can be represented by the linear relation

$$\mathbf{S} = \mathbf{W}^T \Phi + \mathbf{V} \quad (5.1)$$

with Φ as a set of temporal basis functions, $\mathbf{W} = [\mathbf{w}_1, \dots, \mathbf{w}_{N_d}]$ being the spatio-temporal maps (also denoted the regression coefficients) that we are interested in finding, and \mathbf{V} representing source noise. This latter term is actually one of the elements that differentiate the Aquavit model from most of the other spatio-temporal models presented in the literature (see related spatio-temporal work in Section 5.3). The combination of both temporal basis functions and the source noise term ensures that we are not only constrained to a subspace specified by the basis functions but also allows flexibility to search for source activity outside the span of the basis. More specifically our hierarchical model spells

$$\begin{aligned} \mathbf{m}_t &\sim \mathcal{N}(\mathbf{m}_t | \mathbf{A}\mathbf{s}_t, \Sigma_{\mathcal{E}}) \\ \mathbf{s}_t &\sim \mathcal{N}(\mathbf{s}_t | \mathbf{W}^T \boldsymbol{\varphi}_t, \mathbf{D}^{-1}), \quad \mathbf{D} = \text{diag}(\boldsymbol{\alpha}) \\ \mathbf{w}_i &\sim \mathcal{N}(\mathbf{w}_i | \mathbf{0}, \alpha_i^{-1} \boldsymbol{\Lambda}_i^{-1}), \quad \boldsymbol{\Lambda}_i = \text{diag}(\boldsymbol{\lambda}_i). \end{aligned} \quad (5.2)$$

where the covariance of the source noise contribution \mathbf{V} is specified by the hyperparameters $\boldsymbol{\alpha}$ and $\boldsymbol{\varphi}_t$ denotes all basis functions at sample t ($\boldsymbol{\varphi}_t = \Phi_t$). For the spatio-temporal maps (or mixing matrix) \mathbf{W} we assign a zero-mean multivariate Gaussian prior distribution to each \mathbf{w}_i and with precisions $\alpha_i \boldsymbol{\Lambda}_i$. Note that the covariance of the mixing matrix \mathbf{W} is parametrized using $\boldsymbol{\alpha}$ and $\boldsymbol{\Lambda}_i$. We could have omitted $\boldsymbol{\alpha}$ from $p(\mathbf{W})$ without changing the model. However, the present parametrization simplifies the update rules. The prior distribution $p(\mathbf{W})$ enforces the notion of ARD at the level of both space (dipoles) and basis functions such that if any element of $\boldsymbol{\lambda}_i$ is large, the associated basis function at dipole i is deemed irrelevant to explaining the data. During the learning process (see below), many such elements are driven to infinity allowing the sources \mathbf{S} to be represented effectively using only a modest set of basis functions.

For the current sources we apply a multivariate Gaussian distribution with a mean value given by the spatio-temporal maps \mathbf{W} and the temporal basis set Φ . The amount of the source noise contribution \mathbf{V} is controlled by the hyperparameters $\boldsymbol{\alpha}$ with a precision α_i associated to each source. Hereby, we can provide a sparse representation of source noise if present depending on how well the basis functions are specified a priori.

Given the Aquavit model, what we are really interested in is to compute the posterior over the sources. Ideally, we should integrate out all unknowns (including hyperparameters and so forth) and then compute the posterior over the sources $p(\mathbf{S} | \mathbf{Y})$, which contains all possible information about \mathbf{S} conditioned on the observed data \mathbf{Y} . However, the exact posterior in the Aquavit model is computationally intractable similar as with the SOFOMORE model and thus approximation is needed. Here we choose to infer the hyperparameters using the maximum a-posteriori (MAP) point estimate. Given this, the joint posterior distribution over the sources and mixing matrix $p(\mathbf{S}, \mathbf{W} | \mathbf{Y})$ is Gaussian, and

can be computed analytically. Nevertheless, here we also apply a VB [Attias, 2000, Bishop, 2006] approximation $p(\boldsymbol{\theta}|\mathbf{M}) \approx q(\boldsymbol{\theta})$ to speed up the computation. Additionally, we constraint $\boldsymbol{\theta} = \{\mathbf{S}, \mathbf{W}\}$ to be factorized as

$$q(\boldsymbol{\theta}) = q(\mathbf{S}) q(\mathbf{W}) \quad (5.3)$$

with the lower bound given as

$$\mathcal{F} = \langle \log p(\mathbf{Y}, \mathbf{S}, \mathbf{W}) - \log q(\mathbf{S}, \mathbf{W}) \rangle_{q(\mathbf{S})q(\mathbf{W})} \quad (5.4)$$

Following the VB framework described in Section 2.3.1 we iteratively maximize the lower bound of the marginal likelihood w.r.t. the variational distributions $q(\mathbf{S})$ and $q(\mathbf{W})$. This leads to the posterior distributions given in Fig. 5.2.

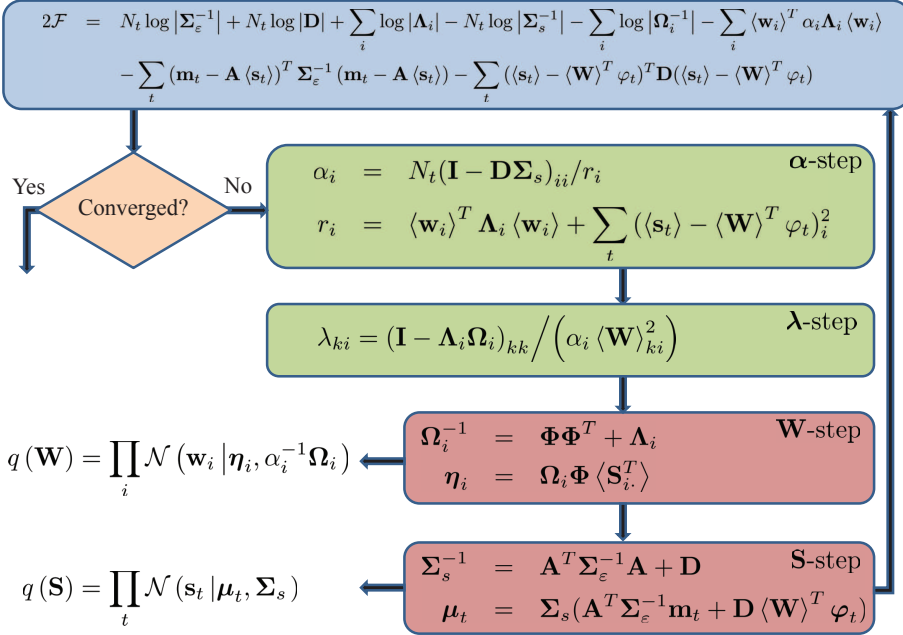


Figure 5.2: **Flow chart of the Aquavit model.** The blue box is the objective function, green boxes are the updates of the hyperparameters, and red boxes are updates associated with the variational posterior distributions $q(\mathbf{W})$ and $q(\mathbf{S})$.

From Fig. 5.2 we note that the hyperparameters are updated as well, since if the variational posterior distributions should be optimal we also need to maximize the hyperparameters. In the VB expected maximization framework the

hyperparameters are updated by taking the derivative of the expected complete data log likelihood w.r.t. the hyperparameters. However, the convergence can be extremely slow and thus for the updates of the hyperparameters in Fig 5.2 we apply Mackay [MacKay, 1992], to derive fixed-point update rules. In practice this leads to significantly faster convergence. Using the variational posterior $q(\mathbf{S})$ and $q(\mathbf{W})$, we obtain the objective function \mathcal{F}

$$\begin{aligned}
2\mathcal{F} = & N_t \log |\Sigma_\varepsilon^{-1}| + N_t \log |\mathbf{D}| + \sum_i \log |\Lambda_i| - N_t \log |\Sigma_s^{-1}| - \sum_i \log |\Omega_i^{-1}| \\
& - \sum_t (\mathbf{m}_t - \mathbf{A} \langle \mathbf{s}_t \rangle)^T \Sigma_\varepsilon^{-1} (\mathbf{m}_t - \mathbf{A} \langle \mathbf{s}_t \rangle) - \sum_i \langle \mathbf{w}_i \rangle^T \alpha_i \Lambda_i \langle \mathbf{w}_i \rangle \\
& - \sum_t (\langle \mathbf{s}_t \rangle - \langle \mathbf{W} \rangle^T \varphi_t)^T \mathbf{D} (\langle \mathbf{s}_t \rangle - \langle \mathbf{W} \rangle^T \varphi_t). \tag{5.5}
\end{aligned}$$

Taking the derivative of \mathcal{F} w.r.t. α_i and λ_{ki} by turn we obtain the updates summarized in Fig. 5.2. For detailed derivations of the variational posterior distributions and the updates for the hyperparameters the reader is referred to consult [Stahlhut et al., 2010a]. However, it should be noted that the notation here and in [Stahlhut et al., 2010a] differs slightly as the notation here have been changed to be consistent with the notation applied throughout the rest of this thesis. In some situations, it is convenient to utilize a single α value to modulate groups of voxels or basis functions, in which case the associated α -update in Fig. 5.2 reduces to an average across the group.

It is worth noting that a special case exists if all elements in \mathbf{W} become zero as the model then resembles the Champagne algorithm [Wipf et al., 2010]. With $\mathbf{W} = \mathbf{0}_{N_k \times N_d}$ the term \mathbf{V} takes over and the sources are modelled as completely independent across time.

5.2 Empirical Evaluation

In order to test the performance of the spatio-temporal Aquavit algorithm we use as benchmark two state-of-the-art methods; the Champagne algorithm [Wipf et al., 2010] and the event sparse penalty (ESP) formulation in [Bolstad et al., 2009]. In Champagne a major difference from Aquavit is that the source time courses are random variables and need to be inferred from data, whereas here we have basis functions which are known in advance. We will denote the latter method as the ESP model. On the other hand the ESP is dependent on obtaining a set of basis functions that are very descriptive of the data.

Even though it could be very interesting to compare the performance of the Aquavit and SOFOMORE methods, they attack the source localization problem from two very different point of views. With the SOFOMORE model focusing on the uncertainty related to the forward fields, the Aquavit attempts to capture the temporal patterns of the current sources. Thus, with the temporal information in this chapter as the primarily focus we devote us to simulations consisting of more challenging temporal patterns than the ones constructed for evaluating the performance of the SOFOMORE model in previous chapter. Of the same reason this chapter will only apply one fixed forward model with no errors in geometry or in the tissue conductivity values.

5.2.1 Random lead field matrix

We first apply the models to a small toy problem where we construct a random generated lead field matrix ($\mathbf{A} : 10 \times 25$) with unit variance. We generate two active sources with correlated times series, as seen in Fig. 5.3. Their times series are created by a sparse set of sine functions. More specifically, we have $s_1(t) = 0.3 \sin(4\pi t) + \sin(60\pi t)$ and $s_2(t) = \sin(10\pi t) + 0.8 \sin(60\pi t)$. Since noise is normally present in electromagnetic recordings, we apply a $\text{SNR} = 1$, with SNR defined as the ratio between the power of the clean signal and the noise.

Our set of temporal basis functions are constructed to include $N_k = 30$ sinusoidals with frequencies between 1 and 30 Hz, denoted Φ^* , with $*$ indicating that the temporal basis functions include the simulated sources. In contrast, we construct a similar temporal basis set Φ , however, now with two of the sine functions (frequency 1 and 5 Hz) used to represent the simulated sources removed. If we compare the simulated sources in Fig. 5.3(a) with the reconstructed sources in Fig. 5.3(b) it is seen that the Champagne algorithm performs quite well. However, taking a closer look it is seen that the estimated temporal signatures by Champagne are corrupted by a number of spikes, due to the modeling of each time point independently of each other. The ESP model on the other hand is very dependent on obtaining an accurate estimate of the true temporal subspace. Given the temporal basis set Φ^* the ESP almost reconstructs the simulated sources perfectly, Fig. 5.3(c). However, examine the ESP performance in a more likely situation, where we cannot be sure that the temporal basis set include all the active sources time series, the source estimates obtained by the ESP model deteriorate quite significantly as seen in Fig. 5.3(d). Due to the fact that the ESP model is restricted to the pre-defined set of temporal basis functions, the model is not able to represent the simulated sources with time series outside the span of the temporal basis set Φ . Indeed, the interpretation that sources both consist of some temporal patterns and ‘source noise’ actually

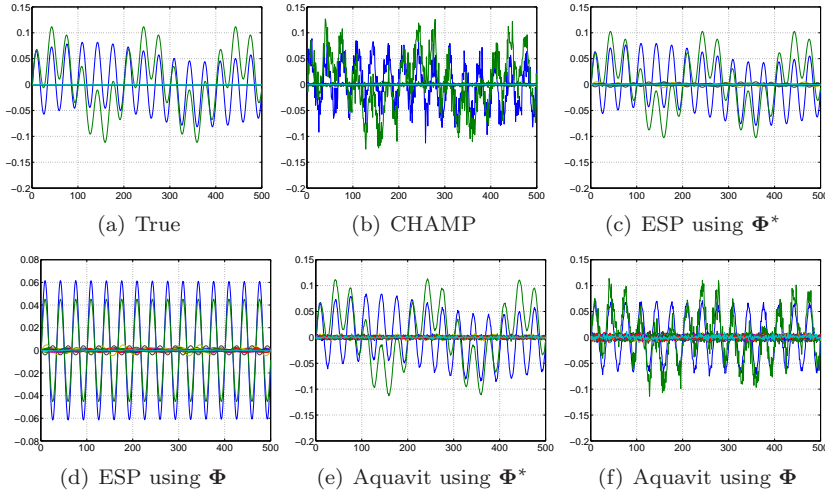


Figure 5.3: **Source estimates** with y-axis representing the amplitude of the sources and the x-axis the time. Φ^* denotes the set of temporal basis functions including the basis functions used to construct the time series of the simulated sources. Mean square error (MSE) on the source estimates: (a) 0, (b) 0.179, (c) 0.011, (d) 0.408, (e) 0.015, and (f) 0.160 .

leads to a balance between these two quantities. Given the true temporal basis set Φ^* the temporal signature of the reconstructed sources are almost perfectly recovered in Fig. 5.3(e). Note how the source noise quantity is taken over in Fig. 5.3(f) and obtains similar patterns as the Champagne algorithm.

5.2.2 Realistic Head Model

To illustrate the applicability of the Aquavit algorithm on more realistic problems we apply the model to a cortical resolution of (~ 500 vertices) obtained by subsampling a more dense set. However, such low resolution or even lower resolution might actually be the case if applied in a context similar to [Bolstad et al., 2009], in which vertices are clustered such that they form a number of patches based on anatomical and functional information. As head model we use a Boundary Element Method (BEM) forward model from SPM8¹. A comparison of the Aquavit algorithm and the ESP model is performed at different SNRs. We construct the simulated current sources in similar fashion as the toy problem above. I.e we generate two artificial current sources from a mixture of sines functions with different frequencies.

¹SPM8 academic software (<http://www.fil.ion.ucl.ac.uk/spm/>).

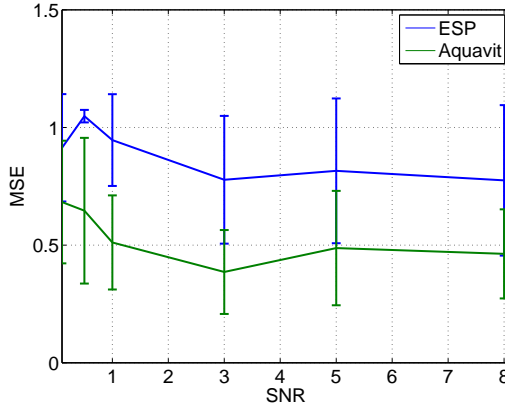


Figure 5.4: Investigation of the **performance of the ESP and Aquavit methods** at different SNR levels and with the number basis functions Φ , $N_k = 30$. At each SNR level 10 experiments were performed and the mean and standard deviation of MSE reported.

Figure 5.5(a) and 5.5(d) show the locations and times series for an example of two simulated sources with the SNR=0.5, respectively. In Fig. 5.5(b) the ESP model localizes the position of source #1 correctly but the temporal signature is not correctly recovered. The Aquavit algorithm on the other hand manages to reconstruct both position and time series of source #1. For source #2 the time series seem better recovered by the ESP model, however with the position of the reconstructed source being misplaced. This misplacement is also reflected in the area under the receiver operating curve (AUC) of 0.541. In contrast Aquavit results in AUC=1, which indicates perfect localization. However, minor deviation in the temporal signature of source #2 exists.

5.3 Related Work

The use of both spatial and temporal information to constrain the source estimates has been advocated for decades. For example, temporal priors have been proposed that operate by penalizing differences in neighboring time points [Baillet and Garnero, 1997], while other priors incorporate temporal smoothness based on second derivatives [Daunizeau et al., 2006]. To describe the temporal dynamics of the sources more flexibly, wavelet temporal basis functions have also received much attention in the M/EEG community. In [Effern et al., 2000] the focus is to represent event related potential (ERPs) with the use of a small set of

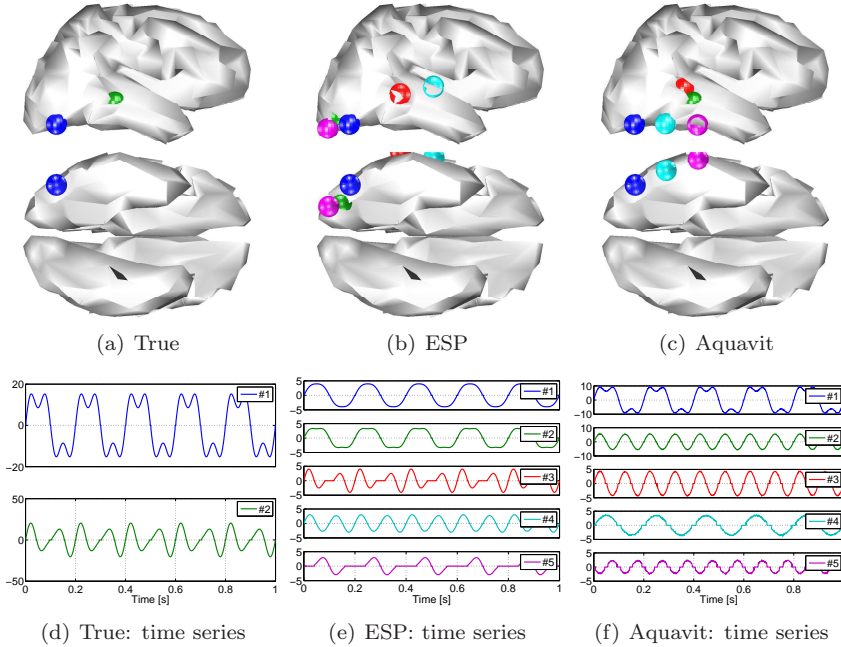


Figure 5.5: **Brain: The five sources with the largest variances.** The sources are color-coded such that their corresponding **time courses** can be seen below the brain images. MSE/AUC measures: (a) 0/1, (b) 0.862/0.541, and (c) 0.403/1.

wavelet bases. Additionally, [Trujillo-Barreto et al., 2008] presents a variational Bayesian approach which tries to represent the M/EEG signal by a small set of coefficients using a wavelet shrinkage procedure.

A quite related method has been presented in [Bolstad et al., 2009] in which an event sparse penalty (ESP) procedure is developed. The approach taken in [Bolstad et al., 2009] seeks a solution composed of a small number of space-time events (STEs) chosen out of a large set of candidates (unlike the method from [Trujillo-Barreto et al., 2008] that operate on a full basis set). In [Bolstad et al., 2009] each STE is a spatio-temporal signal defined in terms of a group of basis functions. Similarly, we have pursued ESI solutions formulated by the sparse spatio-temporal Aquavit algorithm. However, in contrast to [Bolstad et al., 2009], which only computes point estimates by solving a penalized generalized ℓ_1 norm regression problem, the Aquavit consists of a hierarchical Bayesian model such that we can compute a tractable posterior distribution on the unknown sources. Another difference between Aquavit and the work presented in [Bolstad et al., 2009] is that the source generators are restricted to

temporal patterns within a subspace of the set of temporal basis functions used by the method. Likewise, [Ou et al., 2009] apply a spatio-temporal approach in which the measurements and source signals are projected onto temporal basis functions. Thus, the source reconstruction is carried out in the subspace spanned by the temporal basis functions meaning that constructing appropriate temporal basis functions is essential for the method. Projecting the data onto the basis function may be regarded as a preprocessing step in which the data is filtered by a number of different filters corresponding to the basis functions. After filtering the data it can be forwarded to an inverse solver for taking care of the source estimates.

Yet, another study by [Sahani and Nagarajan, 2004] is actually very related to the Aquavit as well. In fact [Sahani and Nagarajan, 2004] also perceives the current sources as a linear combination of a set of temporal functions and a mixing matrix with hyperparameters controlling the relevance of the elements in the mixing matrix. In [Sahani and Nagarajan, 2004] on the other hand the temporal functions are estimated and denoted as pre-sources. Hereby, the method can adapt the temporal information to resemble the temporal patterns observed in the data. This is in contrast to the Aquavit algorithm, which in stead includes a source noise term for capturing source generators not represented by the temporal basis functions.

Finally, there is a study by [Zumer et al., 2008] in which temporal basis functions are learned from the data and afterwards used as fixed basis functions in the source localization process. This work differs in a number of ways from the Aquavit algorithm. One the major differences is how the source are modelled. In [Zumer et al., 2008] the current sources are interpreted as a linear combination of temporal basis functions and a mixing matrix without any hyperparameters facilitating the pruning of elements to zero. Additionally, the approach taken in [Zumer et al., 2008] for modeling source noise contributions differs in the sense that they apply a random distributed mixing matrix for connecting source noise out to the sensors. In contrast Aquavit connects the source noise activity to the sensors through a more anatomical plausible mixing matrix - the forward fields.

5.4 Summary

An overview of the hierarchical structure of the spatio-temporal model Aquavit have been given. The proposed model can be considered as an algorithm balancing the usage of prior information specified by temporal basis functions and modeling the temporal information as samples independent of each other. When no support for the temporal basis functions exists the Aquavit model

models the sources with unsupported temporal patterns as independent across time. A concern with the Aquavit algorithm is the computational burden when the number of temporal basis functions (N_k) increases, as a $N_k \times N_k$ covariance matrix needs to be calculated. However, by choosing an orthonormal basis set the computations is reduced significantly. It would be interesting to see how well the Aquavit algorithm perform on real EEG data when different temporal basis functions are applied.

CHAPTER 6

Source Analysis in Infants

In contrast to the previous chapters this chapter focuses on analyzing real EEG data. While the functionality of the human brain in adults have been extensively examined and yet still far from being fully understood, the functionality of the infant brain is less covered. However, during the last decade this population have gained increasing interest by the neuroimaging community as harmless non-invasive imaging techniques are evolving. In this relation functional imaging by EEG is highly attractive as the measurement can be carried out such that infants can feel comfortable e.g. by sitting on the lap of their parents. To ensure such comfortable setups for infants in other functional imaging modalities like fMRI and MEG are more problematic. In fact studies carried out with fMRI and MEG on infants are primarily with the infants sleeping or sedated, [Altman and Bernal, 2001, Morita et al., 2000, Liu et al., 2008, Cheour et al., 2004]. Requiring the infants to sleep or being sedated significantly limit the types of studies which can be carried out. In this chapter we will examine how pre-verbal infants experience consciousness by analyzing their perception of faces through recording obtained by high-density (128 channels) EEG. The EEG analysis presented in this chapter is carried out in collaboration with Assoc. Prof. Sid Kouider and is currently in preparation, [Kouider et al., 2011, Stahlhut et al., 2011].

6.1 Paradigm

In order to examine face perception during infancy, EEG recordings for three different age groups (5-, 12-, and 15-months-old) were carried out. A schematic illustration of the paradigm is given in Fig. 6.1. The choice of faces as the object of stimuli is far from arbitrary. In the literature it is well known that face-like objects are more attractive for infants than other types of objects as focusing on and decoding faces in infants are used to learn how to identify and to interact with their conspecific [Johnson et al., 1991, Gelskov and Kouider, 2010]. Thus, to maximize the chance of attracting the attention of infants during the experiment a sequence of faces and masked faces is used. While it is well accepted that the reaction time for infants is slower than adults, it is rather unclear how much sensory information that is necessary to provoke a behavioral response in an infant brain, [Gelskov and Kouider, 2010].

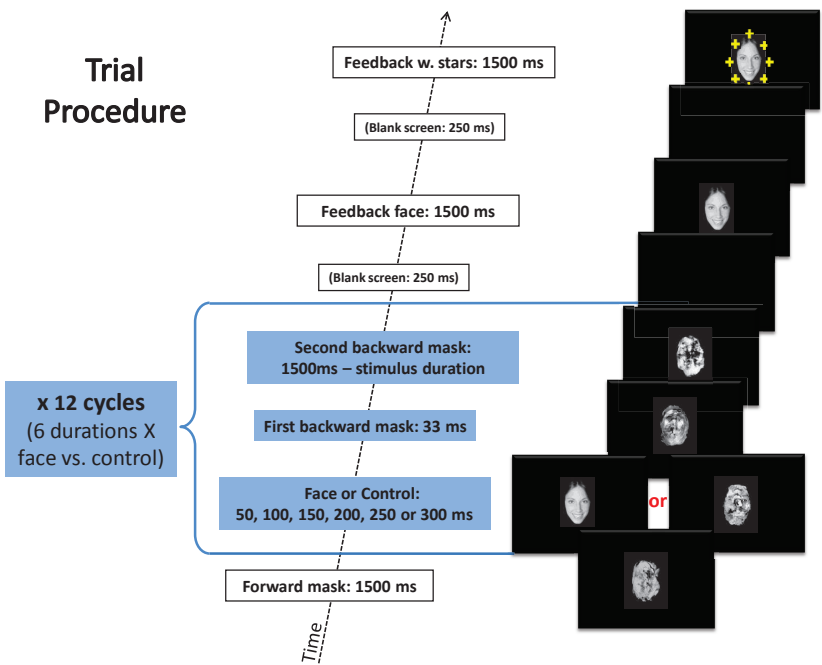


Figure 6.1: **Schematic representation of the paradigm** for investigating face perception in infants. Illustration provided by Sid Kouider and Sofie A.V. Gelskov [Kouider et al., 2011, Gelskov, 2008].

As demonstrated in Fig. 6.1 the paradigm was organized such that infants received either a face or control (scrambled face) image with a duration of either

- 5-months old infants: 50ms, 100ms, 150ms, 200ms, 250ms or 300ms
- 12- and 15-months old infants: 16ms, 33ms, 50ms, 100ms, 150ms, 200ms, or 250ms

Depending on the age group of interest a number of 12 or 14 different stimuli (conditions) were used, as it was expected that 12- and 15-months old infants would be faster to perceive the visual stimuli. A random ordering of the stimuli were used, however, with a restriction that the different types of stimuli had to appear the same amount of times before a similar stimulus could be presented again.

6.2 Method

As EEG source imaging in infants is at its early stage we focus the source analysis in this chapter on standard inverse methods provided by the SPM8 academic software package. First of all due to the fact that the software package is well accepted by the neuroimaging community and the primary goal of this study is to draw researchers attention towards the data set. Second of all this ensures that e.g. psychologists without further programming skills with minimum effort will be able to reproduce or compare their findings with the results presented here and in [Kouider et al., 2011, Stahlhut et al., 2011]. However, as highlighted in Chapter 2-4 a serious concern in general when performing EEG source reconstruction is the choice of forward model as this will influence the final source estimates. Thus, we here apply a trade-off between the complex structure in the SOFOMORE model and the simple standard source reconstruction approaches using a fixed single forward model. The compromise consists of making use of standard inverse method provided by the SPM8 software and performing forward model selection. I.e. we seek the most likely forward model for a given data set at hand among a set of constructed forward models. Of standard source localization approaches we apply are the minimum norm (MN) [Hämäläinen and Ilmoniemi, 1994] and the low resolution electromagnetic tomography (LORETA), [Pascual-Marqui et al., 1994]. A general formulation of the MN and LORETA can be obtained by Tikhonov regularization [Tikhonov and Arsenin, 1977, Hansen, 2010]

$$\mathbf{s}_\lambda = \arg \min \left(\|\mathbf{m} - \mathbf{A}\mathbf{s}\|^2 + \lambda \|\mathbf{H}\mathbf{s}\|^2 \right) \quad (6.1)$$

where the choice of \mathbf{H} differentiates MN and LORETA. In MN \mathbf{H} is an identity matrix whereas in LORETA \mathbf{H} is a Laplacian matrix to ensure spatial smoothness. Of importance is the choice of the regularization parameter λ as this will drive the source solution towards zero for $\lambda = \infty$ and with $\lambda = 0$ no regularization is applied. In SPM8 the implementations of the MN and LORETA methods are in principle extensions of the original MN and LORETA formulations as the SPM implementation applies a Bayesian formulation incorporating temporal priors as well - for further details consult [Friston et al., 2008]. Source estimation is obtained using Restricted Maximum Likelihood (ReML) optimization, originally introduced by [Patterson and Thompson, 1971]. ReML can be formulated such that estimates of the log evidence for a given model can be obtained, [Friston et al., 2007, Friston et al., 2008]. Given the log evidence for a model we can construct a number of forward models and hereby perform model selection in order to obtain the most likely model for describing the observed data. As noted in Chapter 2.3.1 from a Bayesian perspective we are in principle advised not to limit the analysis to model selection but instead we should perform model averaging [Trujillo-Barreto et al., 2004]. Hereby, all models are taken into consideration and we weight their contribution according to the model evidence ratios among all the models of interest. The analysis in this chapter can be regarded as a first attempt for decoding the optimal choice of forward models in EEG using the log evidence as the advisor. Further studies should examine the advantage of Bayesian model averaging when multiple forward models are present.

6.3 Empirical Evaluation

While it is well-known that an infant brain is significantly different from an adult brain e.g. in terms of size, templates of an infant brain are needed when individual MRIs are not present. Since the templates included in the SPM8 software only are appropriate for adults we will make use of templates from the Neurodevelopmental MRI Database¹. MRI templates of infants were kindly provided by Professor John E. Richards². The templates are averages of a set of individual MRI images divided into a number of subgroups throughout ages ranging from newborns of 2 weeks to adults aging 89-years. In the current study we apply the templates of 4.5- and 12-months-old infants [Almli et al., 2007, Richards, 2010a, Richards, 2010b], which are the closest match to the population of interest in this study. We recall from Chapter 3 that a change

¹Neurodevelopmental MRI Database:

<http://jerlab.psych.sc.edu/NeurodevelopmentalMRIDatabase/>

²John E. Richards, Department of Psychology, University of South Carolina, Columbia, SC 29208, USA - <http://jerlab.psych.sc.edu/>

in the conductivity values in adults affects the forward fields and thus influence the source estimates. Naturally, a change of conductivity values for infants will have a similar influence on their source estimates. Thus, we focus the forward model selection on the choices of conductivity ratios between brain and skull. In fact it is believed that the brain:skull conductivity ratio is smaller than the ratio in adults as the skull in infants is thin compared to adults [Reynolds and Richards, 2009]. Consequently, the majority of the forward models that we apply here are constructed with small conductivity ratios. Before we continue to the source reconstruction of how face visibility is processed during infancy, we first validate by simulations that the model evidence indeed can be used as an indicator of what would be an optimal match for the true forward fields that originally mapped the current sources to the sensors.

6.3.1 Simulations

In the validation of the forward model selection in infants use three BEM forward models with different brain:skull conductivity ratios (3, 15, and 45). The true conductivity ratio is 15. We select three different regions on the cortex as illustrated in Fig. 6.2 and assign each of the regions to randomly selected frequency ranges with bandwidth of 5 Hz. The three active regions associated time series can also be found in Fig. 6.2. In the simulations we add white noise to the artificially generated EEG signal such that a SNR of 3 is obtained. Similar to the previous chapters we define SNR as the ratio between the power of the clean EEG signal and the power of the noise.

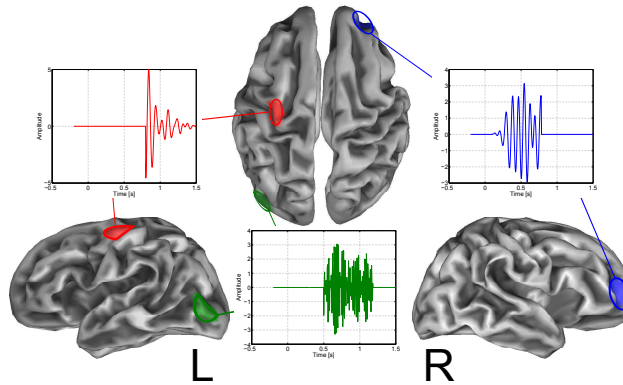


Figure 6.2: The artificially generated source regions to the left and their associated time series to the right.

Fig. 6.3 illustrates the source estimates obtained by the MN method using the three different forward models. Similar, Fig. 6.4 demonstrates the sources estimates obtained by LORETA. We note that both inverse methods visually resemble the true sources in Fig. 6.2 best using the true brain:skull conductivity ratio of 15. More importantly Fig. 6.5(a) reveals that the log evidence measure can be used as an indicator of which forward model to be used. Fig. 6.5(b) illustrates how much variance in the data that is explained by applying the different forward models. We note that the model with a conductivity ratio 45 is the model that leads to the closest variance explained measure to the variance explained of 75.0% when the true source configuration is applied. However, inspecting the source estimates in Fig. 6.3 and 6.4 reveal that source estimates with the conductivity ratio of 45 is in the order of 10^3 times smaller than the true source configuration in Fig. 6.2. In contrast the sources obtained by LORETA and using the forward model with the conductivity ratio of 15, see Fig. 6.4, are very similar to the simulated ones.

6.3.2 Real EEG data

The paradigm illustrated in Fig. 6.1 was carried out with an EEG setup of 128-channel Hydrocel Geodesic Sensor Net from EGI and with an additional physical reference channel at the vertex. The EEG recordings were carried out with a sampling rate of 250 Hz at the Laboratoire de Science Cognitives et Psycholinguistique (LSCP)³. The total number of participants in the different age groups 5-, 12-, and 15-months, were 31, 30, and 30, respectively. We followed a related study with infants, [Bernal et al., 2010], in the preprocessing steps. This has been summarized below:

- Band-pass filtered by 0.5-20Hz.
- Segmented into epochs of 200 ms prior to stimulus onset and 1500 ms post-stimulus.
- Re-reference to an average reference electrode.
- Artifact-rejection: The staff member carrying out the experiment noted bad trials when the infants were not looking at the screen or moving. Additionally, in the remaining ‘good’ trials, channels were examined for artifacts introduced by motions (including eye motion). A maximum of 120 μV local deviation was allowed before being excluded. Also, trials with more than 40% contaminated channels were rejected.

³Laboratoire de Science Cognitives et Psycholinguistique (LSCP), École Normale Supérieure, Paris. Homepage: <http://www.lscp.net/index.php?lang=en>

- Baseline corrected using the 200 ms before onset.
- Averaging the remaining ‘artifact-free’⁴ epochs for each participant within each condition (5-months: 12 conditions, 12- and 15-months: 14 conditions), i.e. we obtain an average response for each condition for each subject.
- Band-pass filtered by 0.5-20Hz.

In order to make a fair comparison of how faces are perceived differently across the age groups, the scalp maps given in Fig. 6.6 only include stimuli durations that all age groups have in common, i.e. 50 ms, 100ms, 150ms, 200ms, and 250ms. The three age groups can be differentiated by the colored boxes surrounding the results. The first two rows within each age group illustrate the mean response to face-stimuli and stimuli of masked-faces, respectively. It is seen that all age groups lead to a quite strong activation of the temporal-occipital regions to both stimuli with faces and masked-faces. For the 12- and 15-months-old infants the visual components seem more localized in time, whereas the occipital-temporal region for the 5-months-old infants is activated for a longer period. Focusing on the mean responses for electrodes located in the occipital-temporal region (see the highlighted electrodes in Fig. 6.7) reveal that the 12- and 15-months indeed lead to a stronger response of shorter duration compared to the 5-months-old infants, c.f. the first column (leftmost column) in Fig. 6.8. The second and third columns in Fig. 6.8 show the responses for each stimuli duration for face and masked faces, respectively. Interestingly, it seems as if the amplitudes of the responses are proportional with the durations of the stimuli. The last column (rightmost) illustrates the difference between face and the control stimuli masked-faces for all durations as well. The difference between the responses to faces and masked-faces reveals the level of visibility of faces for the infants, i.e. we can examine how infants perceive objects as faces and how much time the infants need to recognize that the stimulus is a face and not just a masked-face. A difference between faces and masked-faces appears for the long durations (>100ms) for all age groups. However, in order to determine if there is a significant difference in the responses between faces and masked-faces across subjects we calculate Z-values for the three age groups based on the mean responses to faces and masked-faces for each subject within an age group. The scalp maps of the Z-values are provided in Fig. 6.6 with the rows denoted Z. It is noted that the age groups process faces differently. By comparing the 5-months-old infants with the 12- and 15-months-old infants it is seen that the 5-months-old infants do not have the late component around 1.1s after onset. It is believed that this late component can be associated with cognitive processing such as reflection of what did the image actually illustrate.

⁴Of course epochs cannot be guaranteed to be fully artifact free. However, we can minimize the influence of detectable artifacts.

Source Modeling

In order to get a better understanding of how the infants process faces and masks, source imaging can be used to help revealing these processes by visualizing which brain regions that are active at a given time. The source activity reconstructed in this section is based on a BEM head model using the template MRIs mentioned above. We construct four forward models all with a cortical resolution of 5124 vertices and with the orientations of the forward fields perpendicular to the cortex. The four forward models differ in their conductivities:

- Forward model **A1** with conductivities brain:skull:scalp: 0.33:0.0041:0.33 S/m (conductivities used by most of the EEG toolboxes when analyzing adults).
- Forward model **A2** with conductivities brain:skull:scalp: 0.4:0.3:0.4 S/m.
- Forward model **A3** with conductivities brain:skull:scalp: 0.4:0.04:0.4 S/m.
- Forward model **A4** with conductivities brain:skull:scalp: 0.4:0.02:0.4 S/m.

Source reconstruction using the MN and LORETA methods in SPM8 was performed on the averaged ERP responses across subjects within an age group. Given the four forward models and two inverse methods, we obtain eight different source solutions for each age group. Log evidence and variance explained values are provided in Fig. 6.9-6.11 for the different age groups. According to the model evidence measure we should select forward model **A1** for the 5-months-old infants and the LORETA method. For the 12- and 15-months-old infants **A4** should lead to the best source estimates for describing the observed data. The choice of the forward model for the 5-months-old may seem a bit surprising as it selects a forward model with conductivity values normally applied in studies with adults. However, we need to recall that even conductivity values for adults are an ongoing discussion, and thus care must be taken when comparing the specific values relatively to adults. See e.g. [Oostendorp et al., 2000, Haueisen et al., 1997, Rush and Driscoll, 1969, Homma et al., 1995, Goncalves et al., 2003, Baysal and Haueisen, 2004] for a number of different conductivity values applied in the literature for adults. Also, for the 5-months-old infants the runner-up forward model is **A4**, with a log evidence value relatively close to the log evidence value associated with forward model **A1**. This could be an indication that more conductivity values need to be tested. As an indication of how much variance of the data is described by the different models, variance explained is demonstrated in Fig. 6.9-6.11. It is noted that the forward models providing the largest log evidence (less negative) not necessarily describe most of the observations, due to the fact that noise is also present in the recordings which should not be reconstructed as brain sources.

Given the log evidence measure we can select the source estimates with the largest evidence value for each of the ages and illustrate the source estimates - see Fig. 6.12. Here, we have reconstructed the source activity of the difference between a face response and a masked-face response. Similarly, as with the scalp maps we find one common step for the three age groups around 400ms-600ms, with the ventral (from the primary visual cortex V1 to the inferior temporal lobe) and dorsal (from the primary visual cortex V1 to the parietal lobe) streams being activated. This is well aligned with studies of adults [Bentin et al., 1996], in which the fusiform face area (FFA) has been linked to face processing. We note a second step around 1.0-1.1s after stimuli, mostly apparent in the 12- and 15-months-olds. Here, the temporal lobe is being activated once more. Additionally, the 15-months-old infants have increased activation in the parietal lobe, which are normally associated with spatial processing tasks.

6.4 Related Work

The current study is well aligned with a previous EEG study of face visibility in 5 months old infants [Gelskov and Kouider, 2010]. In this paper it was indicated that the minimum duration of a face to be presented is somewhere between 100-150ms in order for the face to be visible for this specific age group. While the analysis performed in [Gelskov and Kouider, 2010] was carried out in sensor space this chapter have performed an analysis in both sensor and source space. In contrast to [Gelskov and Kouider, 2010] the paradigm presented in this chapter make use of objects centrally placed on a screen in order to minimize movement artifacts in the EEG signal.

The paradigm applied here is quite related to an experiment on adults, which have been presented in [Del Cul et al., 2007]. In [Del Cul et al., 2007] the temporal threshold for adults was found to be in the order of 50ms.

Of alternative EEG work with infants, the [Bernal et al., 2010] can be mentioned. Here, it is demonstrated that 2-year-olds can compute syntactic structures when they listen to a spoken sentence. The discovery that children in fact already at this early stage can compute syntatic structures was performed by EEG source imaging. The analysis of the 2-year-olds applied a single 3-shells model, however without specifying which conductivity values that were used.

Attention of infants have recently been studied using an odd-ball experiment in [Reynolds et al., 2010] with a more detailed description of its source localization approach presented in [Reynolds and Richards, 2009]. Both of these works performed source localization by decomposing the EEG signal into ICA com-

ponents after which each position of the components were reconstructed using an Equivalent Current Dipole (ECD) approach (a parametric method), [Scherg, 1990, Scherg et al., 1999]. In [Reynolds et al., 2010] both 3-spheres and realistic FEM models were applied and a difference in the reconstructed source position was detected.

In contrast to previous work on source localization in infants this work apply a number of forward models and then take advantage of the model evidence measure as an indicator of which forward model is most suitable to describe the observations. In fact the approach of using the model evidence to perform forward model selection (in both infants and adults) is to the authors knowledge the first of its kind in EEG. However, the main idea of using the model evidence as a tool for performing forward model selection have been reported by [Henson et al., 2009] for MEG forward models. As mentioned in Chapter 4.3, [Henson et al., 2009] uses the model evidence to compare different types of head models; single-sphere, overlapping sphere, and BEM. Furthermore, they used the model evidence to examine the cortical resolution and the influence of constraining the source orientation to be perpendicular to the cortex. While forward model selection in this chapter is validated on simulated data and applied to real EEG data, [Henson et al., 2009] focuses entirely on MEG forward model selection in real data. The lack of simulations in [Henson et al., 2009] makes it hard to validate that the model evidence actually can be used as a guide for selecting a forward model that resembles the true forward model best. Additionally, [Henson et al., 2009] does not examine the usage of the model evidence to reveal the optimal conductivity values for the forward model.

6.5 Summary

In this chapter we have analyzed real EEG data of how infants aged 5-, 12-, and 15-months perceive images of faces and masked-faces. As the work presented in this chapter is still in progress, [Kouider et al., 2011, Stahlhut et al., 2011], a first attempt has been to apply standard inverse methods accepted by the neuroimaging community. Indeed, both the SOFOMORE and the Aquavit model would have been highly relevant to apply as well on this data set. Hereby, a comparison of these source estimates with the ones obtained by the minimum norm and LORETA methods implemented in SPM could be obtained.

Of serious concern when studying infants is their level of concentration as movements may be dominating the recordings. As an attempt to neutralize such artifacts in the EEG a staff member conducting the experiment observed the experiment and noted such movements. Additionally, each trial has been exam-

ined for artifacts.

We have proposed to assist the standard inverse methods by forward model selection in order to obtain the most suitable forward model for describing the observed data. Its applicability have been demonstrated on both simulated data and the real EEG data. Forward model selection was performed using the model evidence measure. In fact the model evidence was obtained through a ReML scheme, c.f. [Friston et al., 2007], implemented in SPM8.

The EEG source analysis presented here revealed activation of the ventral and dorsal stream in all three age groups. The activation step arised approximately 400ms after the images were presented. Additionally, an activation step approximately 1000ms after onset was present in the EEG of the 12- and 15-months-old infants. This component was most visibile in the source estimates for the 15-months-old where both temporal and parietal regions were activated. It would be interesting to apply more advanced EEG analysis techniques such as multi-way approaches [Mørup, 2011] in order to take e.g. variability across trials within the same condition into consideration. Multi-way approaches have demonstrated that they are able to capture components that otherwise are averaged out in ‘standard’ EEG analysis.

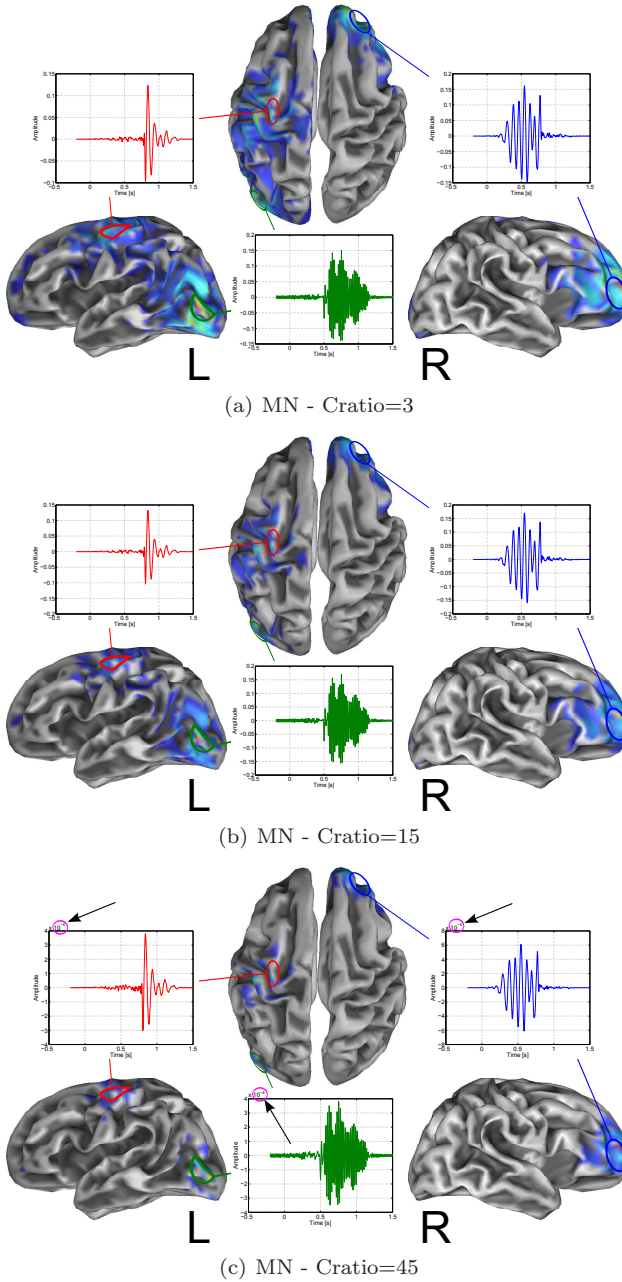
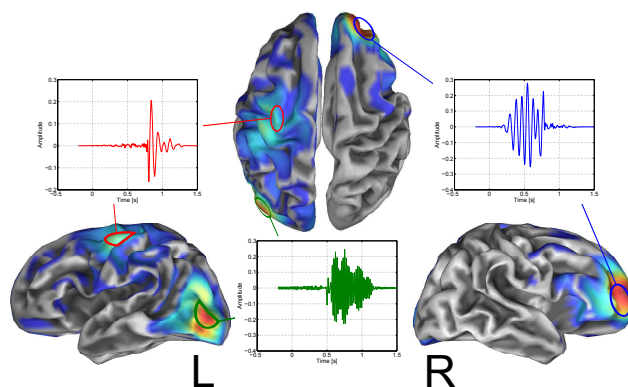
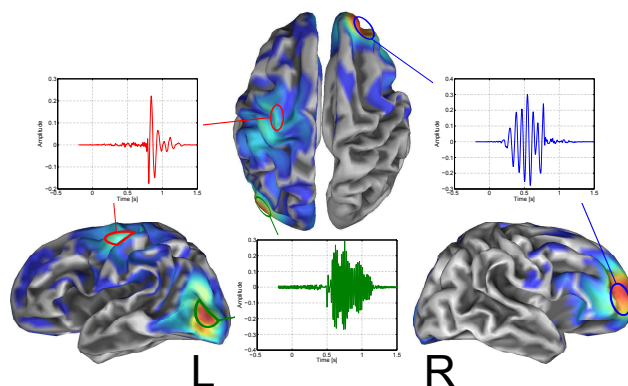


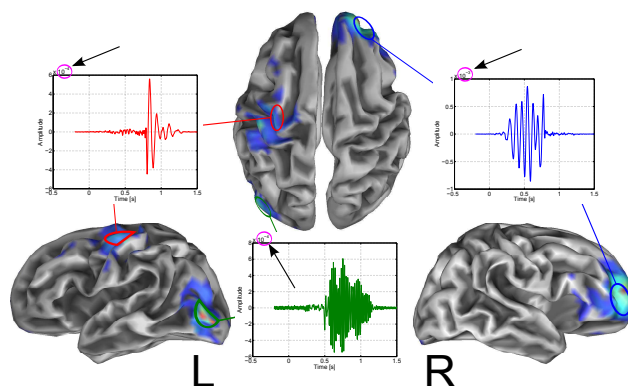
Figure 6.3: **Source estimates obtained by MN using different brain:skull conductivity ratios (Cratio).** Note that the scale of the y-axis is 10^{-4} in (c).



(a) LORETA - Cratio=3



(b) LORETA - Cratio=15



(c) LORETA - Cratio=45

Figure 6.4: **Source estimates obtained by LORETA using different brain:skull conductivity ratios (Cratio).** Note that the scale of the y-axis is 10^{-4} in (c).

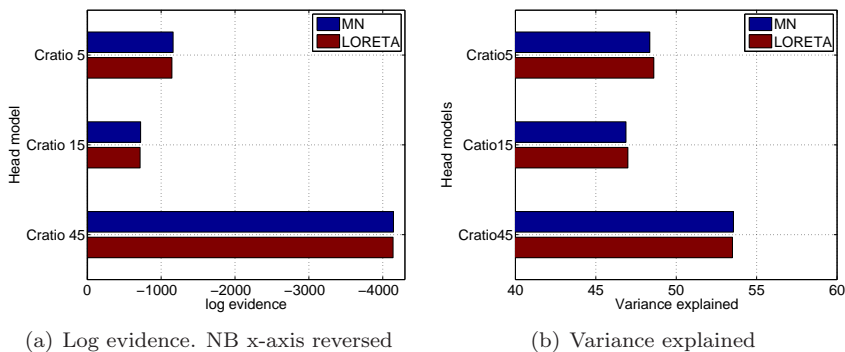


Figure 6.5: **Log evidence values and variance explained** for the two inverse methods MN and LORETA applied to three different forward models with different brain-skull conductivity ratio (Cratio). Applying the true source configuration has a variance explained measure of 75.0

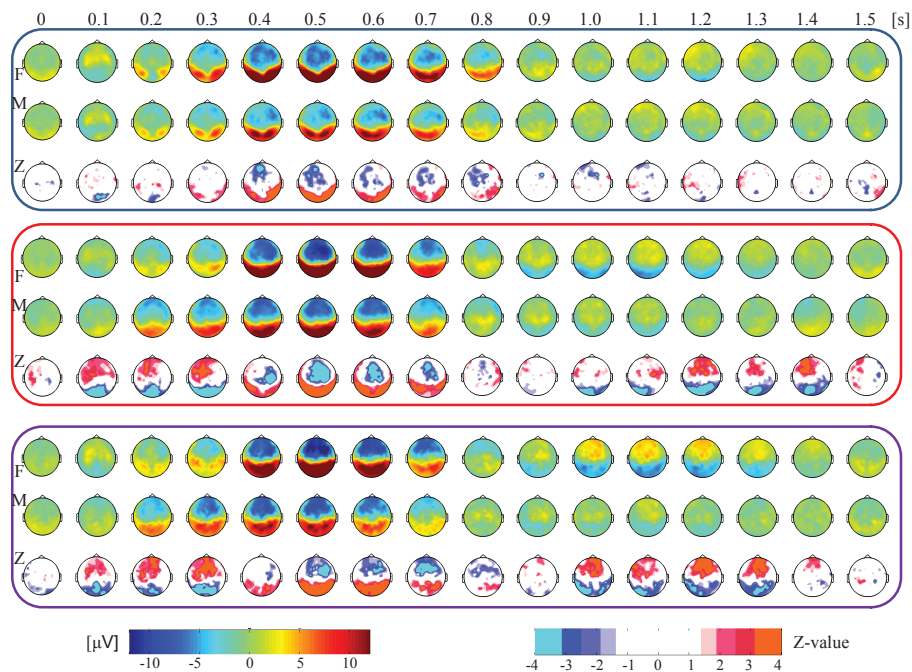


Figure 6.6: **Scalp maps** of the 5- (blue), 12- (red), 15-months-old (purple) infants responses to stimuli faces (F) and masks (M) with steps of 100 ms starting at onset. Scalp maps of the responses to faces and masks are collapsed over the common durations of the age groups (i.e. stimuli durations of 50-250ms). The Z-value of the mean difference of faces and masks for each durations across subjects is given in the rows indicated with Z.

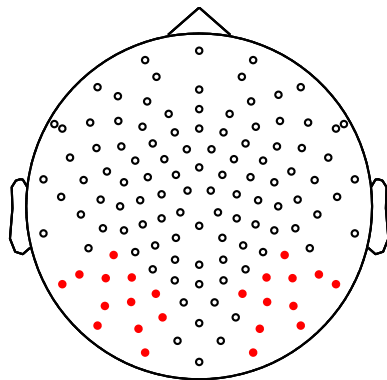


Figure 6.7: **Selected channels in the occipital-temporal region.** Used to explore the response for this region in Fig. 6.8.

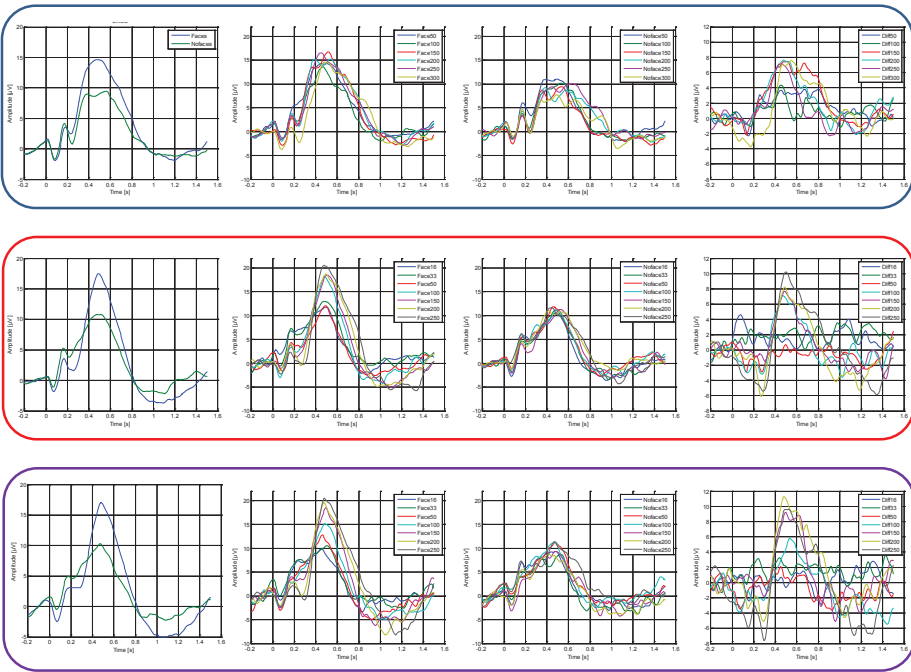


Figure 6.8: **Time series** for mean responses of the channels highlighted in Fig. 6.7 of the 5- (blue), 12- (red), 15-months-old (purple) infants. First column shows the mean responses across the common durations for the age groups (50-250ms).

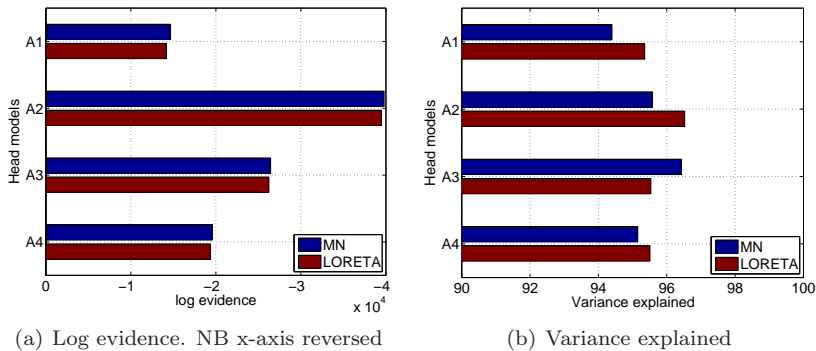


Figure 6.9: **Log evidence values and variance explained** for the two inverse methods MN and LORETA applied to the four forward models **A1**, **A2**, **A3**, and **A4**. Age group **5-months-old** infants.

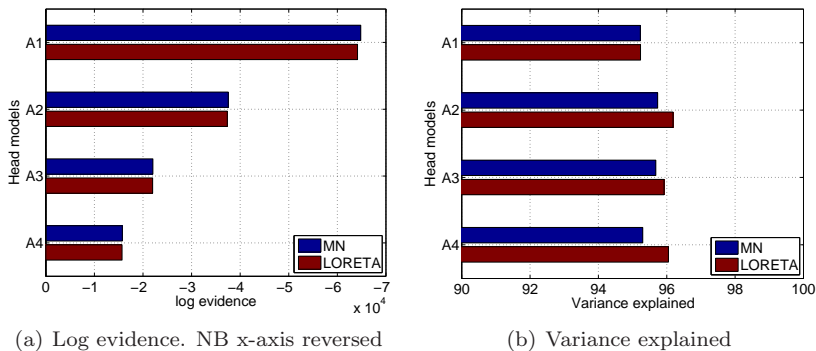


Figure 6.10: **Log evidence values and variance explained** for the two inverse methods MN and LORETA applied to the four forward models **A1**, **A2**, **A3**, and **A4**. Age group **12-months-old** infants.

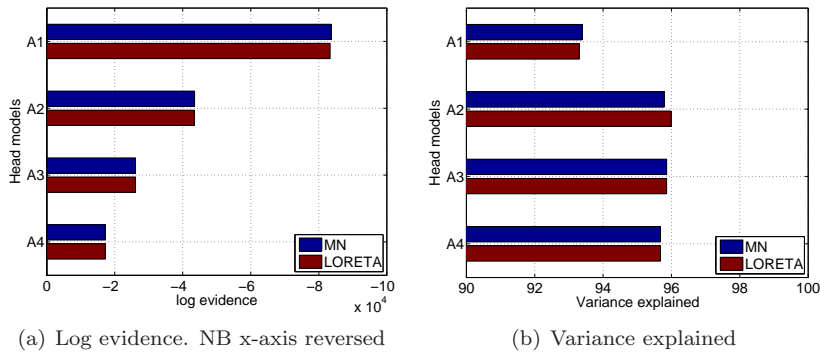


Figure 6.11: **Log evidence and variance explained** for the two inverse methods MN and LORETA applied to the four forward models **A1**, **A2**, **A3**, and **A4**. Age group **15-months-old** infants.

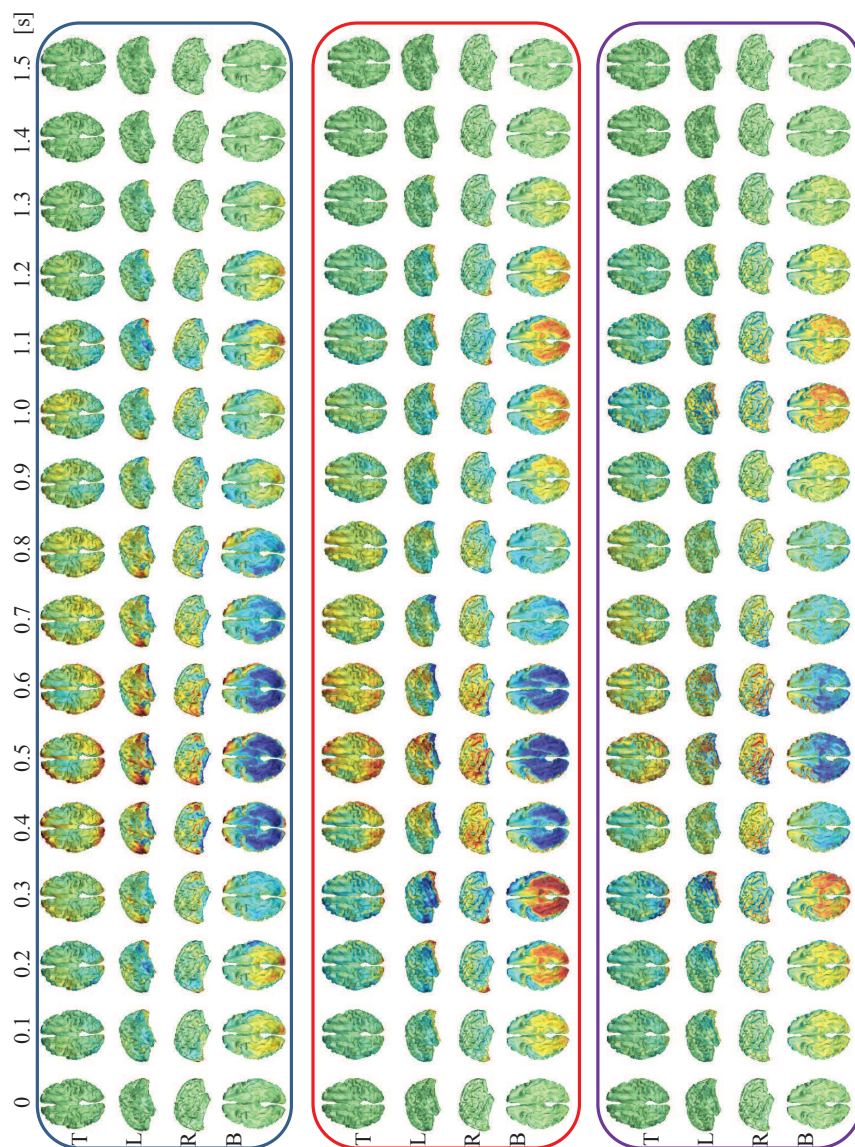


Figure 6.12: **Brain images of reconstructed source activity** of the 5- (blue), 12- (red), 15-months-old (purple) infants. The brain images demonstrate the source activity of the mean difference between the faces and the masks across the common durations (50-250ms) for the age groups. Note that in contrast to the 15-months-old infants, the 5-months-old infants only have very limited activation in the parietal region around 1.0s after onset. T: Top view, L: Left view, R: Right view, and B: Bottom view.

Conclusion

In this thesis we have critically examined the role of the forward fields - connecting current sources within the brain to the sensors at the scalp - in functional brain imaging by EEG. Using EEG as a functional brain imaging tool enables us to obtain a window to the human mind such that we can inspect the temporal processes being carried out. In order to benefit fully from an EEG window to the living brain it must provide us with brain images with high spatial and temporal accuracy. Naturally, the required spatial accuracy is related to the specific brain process needed to be decoded. In this thesis we have investigated how reconstructed sources may be confused with other sources as a consequence of the presence of noise and inadequate forward models. Depending on the amount of source confusion, this spells an erroneous foundation for interpreting a given cognitive task at hand. We have demonstrated that the amount of source confusion is highly dependent on which region a source belong to and the type of forward field errors which are likely to be present. In order to counteract potential errors in the forward model affecting the source estimates, a first attempt of performing simultaneous source and forward model reconstruction (SOFOMORE) have been presented. Among related work [Lew et al., 2007] should be highlighted as this work to the authors knowledge is the first to jointly estimating the unknown brain-skull conductivity ratio in an equivalent dipole setting. The extension by [Plis et al., 2007] to the work presented by [Lew et al., 2007] should be mentioned as well. In [Plis et al., 2007] estimation of the brain-skull conductivity ratio is formulated using a probabilistic distributed

model. However, the approach proposed in our work, SOFOMORE, differs by applying a much more flexible correction of the forward fields represented as a fully stochastic forward model. A serious concern with such an approach is the risk of overfitting. To avoid this we apply an ARD-like prior, which is known for its ability to prune away parameters not contributing significantly to the modeling of the observed data. We have established evidence that SOFOMORE is superior to an equivalent model with a fixed forward model.

While correct forward fields are important for the source estimates, accurate modeling of the relevant temporal processes is just as important when decoding cognitive tasks in human brain. Thus, a novel method making use of temporal basis functions is presented. Applying temporal basis functions is not new to functional neuroimaging, however, in contrast to previous work [Trujillo-Barreto et al., 2008, Bolstad et al., 2009, Sahani and Nagarajan, 2004] we proposed the Aquavit model, a generative hierarchical Bayesian model encouraging sparse spatial and temporal smooth source estimates. The temporal smooth sources are obtained using a sparse set of temporal basis functions. In contrast to, e.g. [Bolstad et al., 2009], the Aquavit model also includes a ‘source noise’ term, which is important when the selected temporal basis set is not suitable for the data at hand. We demonstrated that Aquavit is able to trade-off its temporal representation of the sources between the set of predefined temporal basis functions and the ‘source noise’ to handle situations where the temporal basis set is not completely suitable to describe the data. Interestingly, the Aquavit also demonstrates that in situations where the pre-defined temporal basis set is very suitable for the data, the estimated sources inherit temporal patterns determined by the basis set.

Finally, this thesis has presented an analysis of face visibility in infants. Obtaining accurate forward models for the infants are even more difficult than in adults, as individual MRIs are challenging unless as we want to avoid to sedate the infants. The work included EEG forward model selection based on a free energy approximation to the model evidence. The EEG analysis was carried out using standard inverse methods (minimum norm and LORETA) implemented in the software package SPM8 to ease comparison of results among cognitive psychologists. The analysis demonstrated a temporal-visual component at 400ms after stimulus for all three age groups (5-, 12-, and 15-months-old infants) and for the 12- and 15-months-old infants a later component at 1000ms after stimulus. The component was mostly apparent in the 15-months-old infants and may be associated with complex processing such as reflection of what the image actually illustrated as both temporal and parietal regions were invoked.

Further Directions

During the PhD study a number of exciting research directions have become apparent. In the following suggestions for further work within this field are

listed:

- In continuation of the EEG analysis of data from infants a natural extension of the source localization performed here is to apply the SOFOMORE model to examine its performance and compare it to the results obtained from the forward model selection.
- Combine the SOFOMORE and Aquavit models such that the forward model correction is assisted with temporal reliable functions. This could help restricting the SOFOMORE model. As a first attempt we could assist the SOFOMORE with a fixed temporal basis function without the ‘source noise’ term to reduce the number of unknowns.
- Extension of the Aquavit model: Without too much work this model could be extended to include spatial basis functions as well, which are likely to remove spurious sources.
- Further extension of the Aquavit model: Instead of predefining the temporal basis functions, it could be interesting to estimate the temporal functions jointly with the spatio-temporal maps, i.e. like in an ICA manner just on the source level. However, a concern with such an approach is the risk that the solution might not be identifiable if not assisted with additional constraints such as region specific constraints if this is possible.
- In the current configuration both the SOFOMORE and Aquavit model benefits from ARD-like priors. It would be interesting to apply other sparsity priors such as a Laplace prior and compare the performance with the ARD priors.
- As EEG brain imaging is a highly ill-posed problem, advanced inverse methods may reduce source confusion to some extent, e.g. by jointly optimizing forward fields and the sources, as in the SOFOMORE. However, source confusion still remains a serious concern for EEG analysis. One approach to accomplish ‘trustable’ models is to examine the accuracy of the models in combination with how well the models actually generalize, e.g. by inspecting their predictive accuracy versus their reproducibility. The worth of such techniques has among others been demonstrated in [Strother et al., 2002]. Similarly, a Bayesian formulation based on the same principle was proposed in [Jacobsen et al., 2008]. Such approaches can be used to perform model selection.
- In order to decode the temporal processes carried out by the human brain, so-called functional connectivity analysis [Bressler et al., 2007, Haufe et al., 2009] could be of interest. Hereby, it can be determined how activity occurring in some regions affects other regions in the brain at a later

point. In relation to the source reconstruction presented in this thesis, this would correspond to adding an extra layer of analysis in which the reconstructed sources are inputs and the outputs are directions arrows combining the reconstructed sources in space and time.

- Finally, it could be interesting to explore real-time EEG brain imaging as this would allow new types of experiments to be conducted. For example with online ‘brain state’ detection we would be able to select a proper stimulus depending on the brain state the subject is currently in.

APPENDIX A

SOFOMORE: Combined EEG SOurce and FOrward MOdel REconstruction

Carsten Stahlhut, Morten Mørup, Ole Winther, and Lars K. Hansen. SOFOMORE: Combined EEG SOurce and FOrward MOdel REconstruction. *IEEE International Symposium on Biomedical Imaging: From Nano to Macro (ISBI)*. June 2009.

SOFOMORE: COMBINED EEG SOURCE AND FORWARD MODEL RECONSTRUCTION

Carsten Stahlhut, Morten Mørup, Ole Winther, Lars Kai Hansen

Technical University of Denmark
Department of Informatics and Mathematical Modelling
Richard Petersens Plads, 2800 Lyngby, Denmark.

ABSTRACT

We propose a new EEG source localization method that simultaneously performs SOURCE and FORWARD MODEL REconstruction (SOFOMORE) in a hierarchical Bayesian framework. Reconstruction of the forward model is motivated by the many uncertainties involved in the forward model, including the representation of the cortical surface, conductivity distribution, and electrode positions. We demonstrate in both simulated and real EEG data that reconstruction of the forward model improves localization of the underlying sources.

Index Terms— EEG, source reconstruction, uncertain forward model, hierarchical Bayes, distributed model

1. INTRODUCTION

Functional neuroimaging by PET and fMRI has added considerably to our understanding of the spatial distribution of information processing in the human brain. To better understand the temporal aspects of systems neuroscience electro- and magneto-encephalography (EEG, MEG) are promising modalities due to their high temporal resolution. In addition EEG is attractive as it can be acquired under more naturalistic conditions than PET, fMRI, and MEG. Unfortunately, brain mapping by EEG is an ill-posed problem and relatively strong assumptions are needed for regularization. In the existing literature of source reconstruction significant efforts have been devoted to methods that incorporate spatio-temporal priors to accommodate for focal source distributions under the assumption that the forward model is known. Such inverse methods are typically classified as equivalent current dipole (ECD) or distributed models. ECD models assume that the sources can be described by a small number of ECDs [1], such that the problem is restricted to identify their location and orientation. Distributed models normally assume distributed currents with dipoles oriented perpendicular to the cortical surface. To solve the inverse problem several prior assumptions have been suggested such as mathematical constraints like the (weighted) minimum norm [2] and maximal smoothness [3] properties. Other priors such as anatomical, physiological, and temporal information [4, 5, 6] have also been proposed. Integrating such prior assumptions into the

source reconstruction problem is conveniently implemented in a Bayesian framework. Common to existing source localization methods is that they perform the inversion under the assumption that the forward model is known. However, many noisy processes compromise the forward model, including the representation of the conductivity distribution, the cortical surface, and electrode positions. When 'realistic head models' are constructed from tissue segmentation based on e.g. structural MRI, the geometry of the head model is affected by the resolution and tissue segmentation errors. Such geometric errors can be represented as small magnitude perturbations of the head model shape [7]. Due to the dependence of the head model in the ill-posed EEG source reconstruction problem these uncertainties will also affect the source estimation. Here, we will make a first attempt to correct the uncertain forward model simultaneously with the source localization in order to obtain a more reliable source estimate. Recently, a similar aim was pursued in [8, 9], however, from quite a different perspective than ours. In [8] the geometric structure of the forward model is assumed known, while the skull-brain conductivity ratio is an unknown parameter which is estimated simultaneously with the sources. Similar, [9] reported a probabilistic distributed model to account for uncertainty in the skull conductivity. In this contribution, we suggest a more uncommitted approach in which the complete forward model is considered 'uncertain'. Thus, we approach the EEG source localization problem with a hierarchical Bayesian model that simultaneously perform SOURCE and FORWARD MODEL REconstruction, in short, the SOFOMORE method.

2. METHODS

2.1. The Forward Model

The relationship between the recorded EEG $\mathbf{M} = \{\mathbf{m}_t\}_{t=1}^{N_t}$ from an array of N_e sensors placed on the scalp and the neural current activity in the brain is given by

$$\mathbf{M} = \mathbf{A}\mathbf{S} + \mathbf{E}, \quad (1)$$

where \mathbf{A} is the forward model consisting of a set of forward fields $\{\mathbf{a}_i\}_{i=1}^{N_d}$ corresponding to each of the dipoles in the primary source current denoted $\mathbf{S} = \{\mathbf{s}_t\}_{t=1}^{N_t}$. We assume the

orientations of current dipoles at the vertices to be perpendicular to the cortical surface and the measurement noise \mathbf{E} is modeled as additive zero-mean Gaussian distributed. The forward model was estimated by the SPM5 academic software (<http://www.fil.ion.ucl.ac.uk/spm/>), based on routines from BrainStorm (<http://neuroimage.usc.edu/brainstorm/>) as a basic three-spheres (scalp, skull, brain) head model.

2.2. SOFOMORE: A Hierarchical Bayes Approach

Given the linear relationship in Eq. (1) and the Gaussian noise model, for which we will denote the precision by β , the likelihood, $p(\mathbf{M}|\mathbf{S}, \mathbf{A}, \beta)$, can be expressed as a product of multivariate Gaussian distributions $\mathbf{m}_i \sim \mathcal{N}(\mathbf{A}\mathbf{s}_i, \beta^{-1}\mathbf{I}_{N_e})$. Figure 1 is a graphical representation of the hierarchical model. For the sources we apply a quite similar hierarchical prior as [6], i.e. a zero-mean multivariate Gaussian distribution with a diagonal precision matrix \mathbf{D} , with elements $\alpha = \{\alpha_i\}_{i=1}^{N_d}$, i.e., $\mathbf{s}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{D}^{-1})$. In contrast to [6] we model the uncer-

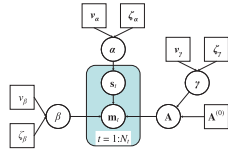


Fig. 1. Graphical representation of hierarchical model that accounts for an uncertain forward model.

tainty of the forward fields, which is performed by independent multivariate Gaussian distributions with prior mean $\mathbf{a}_i^{(0)}$ and precision γ_i , $\mathbf{a}_i \sim \mathcal{N}(\mathbf{a}_i^{(0)}, \gamma_i^{-1}\mathbf{I}_{N_e})$. The prior mean of the i^{th} forward field is obtained from the solution to the forward problem. An assignment of a parameter to each of the forward fields allow us to automatically control which of the forward fields that should be corrected if necessary. Hereby, mainly the forward fields for the active sources have to be corrected and the rest remain unchanged if these precisions are large. Conjugate priors for all the precision parameters have been used, with skewness parameter ν_x and inverse scale parameter ζ_x , see Fig.1, i.e. precision parameter x is modeled with a Gamma distribution $\mathcal{G}(x|\nu_x, \zeta_x)$.

2.2.1. Variational Bayesian Formulation

The Bayesian framework provides the complete joint posterior of sources and parameters given the observed data \mathbf{M} . As a representation of the sources we use the marginal posterior distribution

$$p(\mathbf{S}|\mathbf{M}) = \frac{p(\mathbf{M}, \mathbf{S})}{p(\mathbf{M})} = \frac{\int p(\mathbf{M}, \boldsymbol{\theta}) d\boldsymbol{\theta}_{\setminus \mathbf{S}}}{p(\mathbf{M})}, \quad (2)$$

Table 1. VB updates for hierarchical model with $\langle \cdot \rangle$ denoting the expectation and \mathbf{l}_j the j^{th} lead field i.e. j^{th} row in \mathbf{A} . \mathbf{G} is a diagonal matrix with λ in the diagonal, $\psi(\cdot)$ is the digamma function and its derivative is $\psi'(\cdot)$. Hyperparameters of the form ν_x and ζ_x are updated using Newton-Raphson.

$q(\boldsymbol{\theta})$	VB updates
$\mathcal{N}(\mathbf{s}_i \boldsymbol{\mu}_i, \boldsymbol{\Sigma})$	$\boldsymbol{\mu}_i = \boldsymbol{\Sigma}(\mathbf{A}^T \mathbf{l}_i \boldsymbol{\mu}_i(\beta), \boldsymbol{\Sigma} = ((\beta \mathbf{A}^T \mathbf{A}) + (\mathbf{D}))^{-1}$
$\mathcal{N}(\mathbf{l}_j \boldsymbol{\eta}_j, \boldsymbol{\Psi})$	$\boldsymbol{\eta}_j = \boldsymbol{\Psi} \left((\mathbf{G}) \mathbf{l}_j^{(0)} + (\beta) \sum_{i=1}^{N_t} (\mathbf{s}_i) m_{j,i} \right)$
	$\boldsymbol{\Psi} = \left((\mathbf{G}) + (\beta) \sum_{i=1}^{N_t} (\mathbf{s}_i \mathbf{s}_i^T) \right)^{-1}$
$\mathcal{G}(\alpha_i \hat{\nu}_{\alpha_i}, \hat{\zeta}_{\alpha_i})$	$\hat{\nu}_{\alpha_i} = \nu_{\alpha_i} + \frac{N_t}{2}, \hat{\zeta}_{\alpha_i} = \zeta_{\alpha_i} + \frac{1}{2} \sum_{i=1}^{N_t} (s_{it}^2)$
$\mathcal{G}(\gamma_i \hat{\nu}_{\gamma_i}, \hat{\zeta}_{\gamma_i})$	$\hat{\nu}_{\gamma_i} = \nu_{\gamma_i} + \frac{N_t}{2}, \hat{\zeta}_{\gamma_i} = \zeta_{\gamma_i} + \frac{1}{2} \left\langle (\mathbf{a}_i - \mathbf{a}_i^{(0)})^2 \right\rangle$
$\mathcal{G}(\beta \hat{\nu}_{\beta}, \hat{\zeta}_{\beta})$	$\hat{\nu}_{\beta} = \nu_{\beta} + \frac{N_t N_d}{2}, \hat{\zeta}_{\beta} = \zeta_{\beta} + \frac{1}{2} \sum_{i=1}^{N_t} \left\langle (\mathbf{m}_i - \mathbf{A} \mathbf{s}_i)^2 \right\rangle$
	$\nu_x^{(k+1)} = \nu_x^{(k)} \exp \left\{ -\frac{\psi(\nu_x^{(k)}) - \ln \nu_x^{(k)} + \ln(x) - (\ln x)}{\nu_x^{(k)} \psi'(\nu_x^{(k)}) - 1} \right\}$
	$\zeta_x = \frac{\nu_x}{\hat{\nu}_x}$

where $\boldsymbol{\theta} = \{\mathbf{S}, \mathbf{A}, \boldsymbol{\alpha}, \beta, \boldsymbol{\gamma}\}$, $\boldsymbol{\theta}_{\setminus \mathbf{S}}$ denotes parameter set $\boldsymbol{\theta}$ except for \mathbf{S} , $p(\mathbf{M}, \boldsymbol{\theta})$ is the joint distribution, and $p(\mathbf{M})$ is the marginal likelihood. Due to the hierarchical structure of the model, the marginal likelihood becomes analytically intractable. Thus, approximations are needed and we will apply a standard Variational Bayesian (VB) framework [10], in which the joint posterior $p(\boldsymbol{\theta}|\mathbf{M})$ is approximated by a parameterized simpler distribution $q(\boldsymbol{\theta})$ whose parameters are determined by maximizing the bound on the marginal likelihood obtained through Jensen's inequality,

$$\mathcal{L} = \ln \int p(\mathbf{M}, \boldsymbol{\theta}) d\boldsymbol{\theta} \geq \int q(\boldsymbol{\theta}) \ln \frac{p(\mathbf{M}, \boldsymbol{\theta})}{q(\boldsymbol{\theta})} d\boldsymbol{\theta} \quad (3)$$

The approximate posterior is taken as fully factored in the parameters $\boldsymbol{\theta} = \{\mathbf{S}, \mathbf{A}, \boldsymbol{\alpha}, \beta, \boldsymbol{\gamma}\}$. The resulting sequential VB updates are given in Tab.1¹.

3. EXPERIMENTS

We demonstrate the viability of the SOFOMORE approach on simulated data and real EEG data. We will show that estimation of the forward model matters, by comparing reconstructed sources with the corresponding model with fixed forward model. The results in this section are for clarity obtained by operating directly on the measurements \mathbf{M} without pre-processing of data like bandpass filtering or use of temporal-subspaces. Extensions to include this can be performed and

¹Note that inversion of covariance matrices of size $N_d \times N_d$ appears in the VB update equations, which in the conventional approach can be performed efficiently with the matrix inversion lemma when \mathbf{A} is assumed known. However, due to the inter-dependence of \mathbf{S} and \mathbf{A} the matrix inversion lemma cannot be used efficiently. Thus, in the implementation we have performed the VB updates in a subspace spanned by the basis of the lead fields (rows in \mathbf{A}).

will in general improve performance, future work will concern this important issue.

3.1. Simulations

In the simulations we construct a source signal of a half sine of duration 50ms with a starting time at $t=25$ ms for a small set of sources located in the occipital lobe, as illustrated in the SPM glass-brain representation in Fig.2(a). Note that minor activity seems to appear at the inner part of the left hemisphere, which is due to the mapping from cortex to the glass-brain representation. The simulated clean EEG was constructed with forward model consisting of 3-spheres with tissue conductivities brain:skull:scalp = 0.33:0.0041:0.33S/m (ratio 1:1/80:1) and dimension 128×7204 . We corrupted the clean EEG with 'realistic EEG noise' obtained from the pre-stimulus period in an evoked EEG study described in Sec. 3.2. We applied a signal-to-noise ratio $\text{SNR}=10$, with $\text{SNR} = P_{\text{EEG}}/P_{\text{noise}}$, where P_{EEG} and P_{noise} are the power of the clean EEG and noise, respectively. Conductivity errors were added to one of the forward models used in the source localization. In the source reconstruction we used a cortical resolution of 4004 vertices, such that the 'true' forward model has a higher spatial resolution. Figures

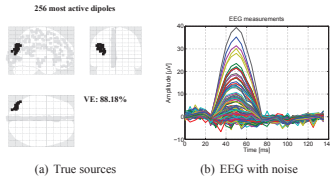


Fig. 2. Simulated source density and EEG. VE: Variance explained.

3(a) and 3(b) show the estimated source densities at the time $t=50$ ms, which corresponds to the time point where we have the maximum amplitude of the EEG. In these figures no corrections of the forward fields are performed but different conductivities have been used. In both cases erroneous activity is estimated in the right pre-motor area (indicated with a circle in the figure). However, a 'correct' choice of conductivities also leads to quite good localization of the true sources. In contrast an erroneous choice of conductivities in Fig.3(a) results in poorer reconstruction of the true sources. The poorer reconstruction can also be seen from a lower log-evidence value and an increase in the variance explained (VE) compared to the correct one in Fig.2(a). The percentages of variance explained in both models are larger than the 'true' value (88.2%), indicating some overfitting. In Figs.3(c)-(d) we use the same forward models as in Figs.3(a)-

(b) as prior means ($\mathbf{A}^{(0)}$) in the SOFOMORE model. Indeed the combined Bayesian estimation of forward model and source density leads to a better estimation of the true source density in both cases. Moreover, the resulting estimates are quite similar. However, the model with 'correct' prior mean explains 91.7% of the data variance, which is closer to the true value than the one with the erroneous forward model taken as prior mean. Additionally, the model in Fig. 3(d) has a much larger log-evidence value, indicating a more likely model.

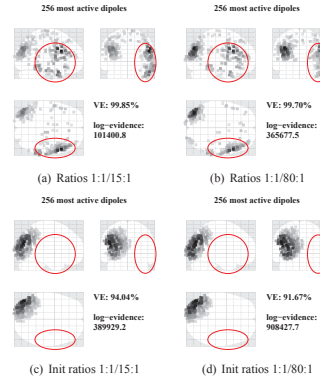


Fig. 3. Estimated activity at $t = 50$ ms and weighted by the square root of the inverse variance α (i.e. $s_{it}/\sqrt{\alpha_i}$). First row: No corrections of forward fields. Second row: Corrections are integrated into the source localization method. Different conductivity ratios are used. Figs.(a)+(b) misleading activity inside circles - compare with Fig.2(a). VE: Variance explained.

3.2. Real EEG data

This EEG data set is part of a multi-modal study on face perception, where faces and scrambled faces were presented for 600ms every 3600ms to a subject. A detailed description of the experiment is given in [11] and <http://www.fil.ion.ucl.ac.uk/spm> (where the data is available for download). In this contribution we reconstruct the average event related potential (ERP) of trials involving real faces as stimuli. In Figs. 4(a) and 4(b) the estimated activity is illustrated at $t=170$ ms after stimulus, without and with integration of the reconstruction of the forward fields respectively. When there is no correction of the forward fields the reconstruction results in quite scattered activity, with a prominent

voxel located in the right temporal lobe and less prominent voxels in the right occipital region. In contrast the SOFOMORE model leads to improved localization of activity in the visual cortex with minor activity in the fusiform gyrus, which is known to be connected with face processing. This result is well-aligned with results reported in [12]. Please note that no spatial smoothing has been performed in our results, which would probably give better localized sources for the case without corrections of \mathbf{A} . Moreover, the columns of the forward matrix can be normalized such that the enhancement of the superficial sources is reduced.

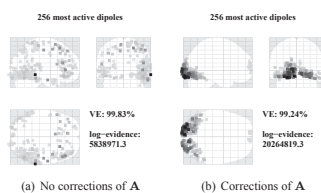


Fig. 4. Estimated activity $t = 170\text{ms}$ after stimulus and weighted by the square root of the inverse variance α when no corrections of forward fields are performed and when corrections are integrated into the source localization method. Tissue conductivities brain:skull:scalp = 0.33:0.0041:0.335/m are used. VE: Variance explained.

4. CONCLUSION

We proposed a hierarchical Bayesian framework for simultaneous source and forward model reconstruction aimed at a improved source density estimate. Results from simulations and real EEG data illustrated the applicability of the model. We found that the combined reconstruction resulted in more localized activity with fewer large distance errors, in comparison with a similar model with a fixed forward model. Overfitting is a serious concern in the extended model. By invoking a flexible prior on the forward model corrections relative to the prior mean - the standard forward model - are only made where it is necessary, hence reducing overfit. Future work concerns the performance of the SOFOMORE model on more realistic head models based on boundary element method and finite element method and more realistic priors with temporal structure.

5. REFERENCES

[1] J. C. Mosher, P. S. Lewis, and R. M. Leahy, "Multiple dipole modeling and localization from spatio-temporal

meg data," *IEEE Trans. on Biomed. Eng.*, vol. 39, no. 6, pp. 541–557, 1992.

- [2] M.S. Hämäläinen and R.J. Ilmoniemi, "Interpreting magnetic fields of the brain: minimum norm estimates," *Med. Biol. Eng. Comput.*, vol. 32, pp. 35–42, 1994.
- [3] R. D. Pascual-Marqui, C. M. Michel, and D. Lehmann, "Low resolution electromagnetic tomography: a new method for localizing electrical activity in the brain," *Int J. Psychophysiol.*, vol. 18, pp. 49–65, 1994.
- [4] S. Baillet and L. Garnero, "A bayesian approach to introducing anatomo-functional priors in the eeg/meg inverse problem," *IEEE Trans. on Biomed. Eng.*, vol. 44, no. 5, pp. 374–385, 1997.
- [5] C. Phillips, M.D. Rugg, and K.J. Friston, "Anatomically Informed Basis Functions for EEG Source Localization: Combining Functional and Anatomical Constraints," *NeuroImage*, vol. 16, no. 3, pp. 678–695, 2002.
- [6] M.-A. Sato, T. Yoshioka, S. Kajihara, K. Toyama, N. Goda, K. Doya, and M. Kawato, "Hierarchical bayesian estimation for MEG inverse problem," *Neuroimage*, vol. 23, no. 3, pp. 806–26, November 2004.
- [7] N. von Ellenrieder, C.H. Muravchik, and A. Nehorai, "Effects of geometric head model perturbation on the EEG forward and inverse problems," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 3, pp. 421–429, March 2006.
- [8] S. Lew, C. Wolters, A. Anwander, S. Makeig, and R.S. MacLeod, "Low resolution conductivity estimation to improve source localization," in *New Frontiers in Biomag. Proceedings of the 15th Int. Conf. on Biomag.*, 2007, vol. 1300 of *Int. Congress Series*, pp. 149–152.
- [9] S.M. Plis, J.S. George, S.C. Jun, D.M. Ranken, P.L. Volegov, and D.M. Schmidt, "Probabilistic forward model for electroencephalography source analysis," *Physics in Medicine and Biology*, vol. 52, no. 17, pp. 5309–5328, 2007.
- [10] C.M. Bishop, *Pattern Recognition and Machine Learning*, Springer, NY 10013 (USA), 2006.
- [11] R.N.A. Henson, Y. Goshen-Gottstein, T. Ganel, L.J. Otten, A. Quayle, and M.D. Rugg, "Electrophysiological and hemodynamic correlates of face perception, recognition and priming," *Cerebral Cortex*, vol. 13, pp. 793–805, 2003.
- [12] K. Friston, L. Harrison, J. Daunizeau, S. Kiebel, C. Phillips, N. Trujillo-Barreto, R. Henson, G. Flandin, and J. Mattout, "Multiple sparse priors for the M/EEG inverse problem," *NeuroImage*, vol. 39, pp. 1104–1120, 2008.

APPENDIX B

Hierarchical Bayesian Model for simultaneous EEG Source and Forward Model Reconstruction (SOFOMORE)

Carsten Stahlhut, Morten Mørup, Ole Winther, and Lars K. Hansen. Hierarchical Bayesian Model for simultaneous EEG Source and Forward Model Reconstruction (SOFOMORE). *IEEE International Workshop on Machine Learning for Signal Processing (MLSP)*. September 2009.

HIERARCHICAL BAYESIAN MODEL FOR SIMULTANEOUS EEG SOURCE AND FORWARD MODEL RECONSTRUCTION (SOFOMORE)

Carsten Stahlhut, Morten Mørup, Ole Winther, Lars Kai Hansen

Technical University of Denmark
Department of Informatics and Mathematical Modelling
Richard Petersens Plads, DK-2800 Kgs. Lyngby, Denmark.

ABSTRACT

In this paper we propose an approach to handle forward model uncertainty for EEG source reconstruction. A stochastic forward model is motivated by the many uncertain contributions that form the forward propagation model including the tissue conductivity distribution, the cortical surface, and electrode positions. We first present a hierarchical Bayesian framework for EEG source localization that jointly performs source and forward model reconstruction (SOFOMORE). Secondly, we evaluate the SOFOMORE model by comparison with source reconstruction methods that use fixed forward models. Simulated and real EEG data demonstrate that invoking a stochastic forward model leads to improved source estimates.

1. INTRODUCTION

Electroencephalography (EEG) measures the electrical activity that arises from neurons in the brain by an array of sensors placed on the scalp. We are interested in reconstruction of the EEG source generators, which is known as the inverse problem with a mapping from source space to sensor space that is many to one. Despite the fact that EEG source reconstruction is a severe ill-posed problem it contains highly interesting information for the functional imaging community due to its high temporal resolution in comparison with functional magnetic resonance imaging (fMRI) and positron emission tomography (PET). The relation between the measured EEG signal and the current sources within the brain can under the quasi-static approximation of Maxwell's equation be expressed as a linear combination of the sources in the so-called forward problem. Given the measured EEG signal $\mathbf{M} \in \mathbb{R}^{N_s \times N_t}$ and the current sources $\mathbf{S} \in \mathbb{R}^{N_d \times N_t}$, with N_s , N_d , and N_t denoting the number of channels, dipoles, and time samples, respectively, the forward relation is given by, [1],

$$\mathbf{M} = \mathbf{A}\mathbf{S} + \mathcal{E}, \quad (1)$$

where the noise \mathcal{E} is assumed additive. The interrelationship between the sensors and the current sources is given

by the lead field matrix/forward model $\mathbf{A} \in \mathbb{R}^{N_s \times N_d}$ with the rows referred to as the lead fields for the sensors and the columns as the forward fields for the sources. In this contribution we assume the orientations of the dipoles at the vertices to be perpendicular to cortex. However, a more flexible orientation can be incorporated by inclusion of three columns in \mathbf{A} per dipole corresponding to the three directional components. Different levels of complexity of the head model exist, where the spherical head model, the boundary element method (BEM), and finite element methods (FEM) are examples of increasingly complex models, [2], [3].

To obtain a unique solution when solving the inverse EEG problem additional information or constraints such as anatomical, physiological, and mathematical properties are needed, [4]–[7]. This has led to a development of a vast number of inverse methods in the literature. One approach used is to restrict the estimated current sources to a limited number of current dipoles. This can lead to an overdetermined inverse problem where the limited set of dipoles is then fitted to the data by minimizing a measure of reconstruction error such that the location, orientation, and amplitude of the dipoles can be obtained, [8]. A different class of inverse methods is the so-called distributed models also known as source imaging techniques, where a distribution of the current sources is recovered. The well-known minimum norm (MN) estimate [9] reconstructs a source distribution that minimizes the l_2 -norm of the measurements and the explained EEG by a given source configuration.

Embedding prior assumptions of the sources \mathbf{S} into the reconstruction problem is conveniently implemented in a Bayesian framework with the use of the prior distribution $p(\mathbf{S})$. In fact this important choice of prior primarily differentiates the different source localization methods. Given the observed data we can now use the marginal posterior distribution as representation of the unknown sources,

$$p(\mathbf{S}|\mathbf{M}) = \frac{p(\mathbf{M}, \mathbf{S})}{p(\mathbf{M})} = \frac{\int p(\mathbf{M}|\boldsymbol{\theta}) p(\boldsymbol{\theta}) d\boldsymbol{\theta}_{\mathbf{S}}}{p(\mathbf{M})} \quad (2)$$

where $\boldsymbol{\theta}$ is a set of parameters for the model including \mathbf{S} and $\boldsymbol{\theta}_{\mathbf{S}}$ denotes the parameter set $\boldsymbol{\theta}$ except for \mathbf{S} . To obtain the

posterior distribution for \mathbf{S} we marginalize all other parameters than \mathbf{S} . The distribution $p(\mathbf{M})$ is the marginal likelihood also referred to as the model evidence and is given by $p(\mathbf{M}) = \int p(\mathbf{M}|\boldsymbol{\theta}) p(\boldsymbol{\theta}) d\boldsymbol{\theta}$. The model evidence is important for model comparison. Since all parameters are integrated out to obtain the model evidence, the Bayesian framework naturally favors simpler models over more complex models that a priori can model a larger range of data sets. Unfortunately, the marginal likelihood becomes intractable for most prior distributions, and thereby the posterior $p(\mathbf{S}|\mathbf{M})$ cannot be computed either. Thus, approximate methods e.g. Laplace approximation [10], variational Bayesian (VB) [6], [11], and markov chain monte carlo (MCMC) sampling [12] have been used.

While existing source localization methods perform the inversion under the assumption that the forward model is known, we will treat the source reconstruction problem without making this assumption about the forward model. This is motivated by the many noise processes that contribute to the forward model, including the representation of the cortical surface, the conductivity distribution, and electrode positions. When ‘realistic head models’ are constructed from tissue segmentation based on e.g. structural MRI, the geometry of the head model is affected by the resolution and tissue segmentation errors. The issue of modeling the forward process was recently pursued in [12], [13] from quite a different viewpoint than ours. In [13] the basic structure of the forward model is assumed known, while the skull-brain conductivity ratio is an unknown parameter which is estimated simultaneously with a few sources in a dipole fit setting. Similarly, [12] has proposed a probabilistic approach for distributed models to account for uncertainties in the skull conductivity. Here, we apply a more unconstrained approach by modeling the whole forward model as uncertain using a prior distribution for the forward model. We propose a first attempt for distributed models to perform simultaneous source and forward model reconstruction, in short the SOFOMORE model [14]. In this paper we focus on the evaluation of the performance of modeling the forward propagation model in the SOFOMORE model in relation to the MN method.

2. METHODS

Given the linear relation in Eq. 1 and if we assume the noise to be time independent multivariate Gaussian distributed, the likelihood for a single time point t can be expressed as $p(\mathbf{m}_t|\mathbf{s}_t, \boldsymbol{\Sigma}_{\mathcal{E}}) = \mathcal{N}(\mathbf{m}_t|\mathbf{A}\mathbf{s}_t, \boldsymbol{\Sigma}_{\mathcal{E}})$ where $\boldsymbol{\Sigma}_{\mathcal{E}}$ is the noise covariance matrix. For simplicity we assume no temporal correlation. However, this can also be integrated as outlined in [10]. In the remainder of this paper we assume $\boldsymbol{\Sigma}_{\mathcal{E}} = \beta^{-1}\mathbf{I}_{N_c}$. In a minimum norm setting a multivariate Gaussian prior for the sources with zero mean and covari-

ance $\alpha^{-1}\mathbf{I}_{N_d}$ is assumed. Moreover, it is assumed that the forward propagation model $\mathbf{A} = \mathbf{A}^{(0)}$ is known. With the use of Bayes rule Eq. 2 it is seen that the posterior distribution is maximized by

$$\boldsymbol{\Sigma}_{\mathbf{M}} = \left(\mathbf{A}^{(0)} (\alpha^{-1}\mathbf{I}_{N_d}) \mathbf{A}^{(0)T} + \beta^{-1}\mathbf{I}_{N_c} \right)^{-1} \quad (3)$$

$$\mathbf{S}_{\text{MN}} = (\alpha^{-1}\mathbf{I}_{N_d}) \mathbf{A}^{(0)T} \boldsymbol{\Sigma}_{\mathbf{M}} \mathbf{M}. \quad (4)$$

It is noted that since the likelihood and prior are both Gaussian distributions, the posterior and marginal likelihood will also be Gaussian distributions. The estimation of the sources, the precision parameters α and β are performed using a standard expectation-maximization (EM) scheme [11].

In contrast to the MN formulation we here propose a hierarchical model that incorporates corrections of the forward fields simultaneously with the source estimation. As prior for the current sources we use a zero-mean multivariate Gaussian distribution with a diagonal precision matrix \mathbf{D} , where the diagonal elements are $\alpha = \{\alpha_i\}_{i=1}^{N_d}$, $\mathbf{s}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{D}^{-1})$. Hereby, the dipoles can have different variances corresponding to some dipoles being expected to be more active than others. The current variances are assumed to not change over the time period $t = 1 : N_t$ and are estimated by an automatic relevance determination prior (ARD) quite similar to [6]. Thus, we have $p(\alpha_i|\nu_{\alpha_i}, \zeta_{\alpha_i}) = \mathcal{G}(\alpha_i|\nu_{\alpha_i}, \zeta_{\alpha_i})$ where $\mathcal{G}(\cdot)$ is the Gamma distribution with skewness parameter ν_{α_i} and inverse scale parameter ζ_{α_i} , such that the mean is given by $\nu_{\alpha_i}/\zeta_{\alpha_i}$ and the variance $\nu_{\alpha_i}/\zeta_{\alpha_i}^2$.

We regard the forward model as a stochastic process, in which the forward propagation model $\mathbf{A}^{(0)}$ is used as a mean value in a multivariate Gaussian prior. Each of the forward fields are modeled as independent, such that the prior distribution of a single forward field is given by prior mean $\mathbf{a}_i^{(0)}$ (i ’th column in $\mathbf{A}^{(0)}$) and precision γ_i , i.e.

$$p(\mathbf{A}|\mathbf{A}^{(0)}, \boldsymbol{\gamma}) = \prod_{i=1}^{N_d} \mathcal{N}(\mathbf{a}_i|\mathbf{a}_i^{(0)}, \gamma_i^{-1}\mathbf{I}_{N_c}) \quad (5)$$

where $\boldsymbol{\gamma} = \{\gamma_i\}_{i=1}^{N_d}$. Assigning each of the forward fields a precision parameter allows us to differentiate between the amounts of correction of the forward fields that should be performed. This is motivated by the belief that some regions in the brain are more affected than others. Similar to the precision parameters of the current sources, the precision parameters for the forward fields are also modeled by conjugate prior distributions, i.e. $\gamma_i \sim \mathcal{G}(\gamma_i|\nu_{\gamma_i}, \zeta_{\gamma_i})$. Likewise, the inverse noise variance is assumed to be Gamma distributed, $\mathcal{G}(\beta|\nu_{\beta}, \zeta_{\beta})$. We note that the MN estimate is a special case of the SOFOMORE solution, where the prior for the forward fields are just delta functions centered at $\mathbf{a}_i^{(0)}$ and moreover the values of precisions for the current sources are all equal.

Given the hierarchical structure of the SOFOMORE model with the parameters $\theta = \{S, A, \alpha, \gamma, \beta\}$ the marginal posterior distribution of the current sources becomes analytically intractable, thus, approximations are needed. We apply a standard VB framework [11], in which a parameterized simpler distribution $q(\theta)$ approximates the true joint posterior $p(\theta|M)$. In the VB framework the parameters of the $q(\theta)$ distribution are determined with a maximization of a lower bound of the marginal likelihood obtained by the Jensen's inequality,

$$\mathcal{L} = \ln \int p(M|\theta) p(\theta) d\theta \geq \int q(\theta) \ln \frac{p(M|\theta) p(\theta)}{q(\theta)} d\theta \quad (6)$$

We assume factorization in the parameters $\theta = \{S, A, \alpha, \gamma, \beta\}$ of the approximate posterior $q(\theta)$. This leads to the sequential VB updates as illustrated in Fig. 1, where each of the approximated marginal posterior distributions of the parameters can be seen.

3. EXPERIMENTS

We demonstrate the effects of the SOFOMORE model source reconstruction for both simulated and real EEG data. We present the recovered source density estimates both with and without estimated forward fields. Besides the MN method with a fixed forward model, we also compare the SOFOMORE model with a model with a similar hierarchical structure as the SOFOMORE, however, with a fixed forward field, i.e. no A- or γ -steps are performed. We denote this model as the ARD model, due to its ARD prior on the sources. For clarity of the results, the methods work directly on the measurements M , i.e. no pre-processing is performed. However, pre-processing will in general improve the performance. As validation metrics we use the mean square error (MSE), variance explained (VE), degree of focalization (DF), and area under the receiver operating curves (AUC). Definition of the validation metrics is given in Tab. 1 in section 3.1.

3.1. Simulations

In the simulations a small cortical area in the left occipital lobe is simulated as active. The source signal consist of a half sine of duration 50ms starting at $t=25$ ms. The simulated sources at $t=50$ ms are shown in the SPM glass-brain representation in Fig. 2. Due to the mapping from cortex to the glass-brain representation, minor activity seems to appear at the inner part of left hemisphere. Moreover, Fig. 2 shows the time series of the simulated sources. The first 256 rows in the image are the 256 most active sources and the last 256 rows below the black horizontal line in the image are 256 randomly drawn sources with the 256 most active sources excluded. This will give an idea of the

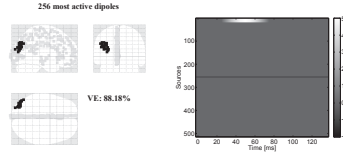


Fig. 2. Left: Simulated sources at $t=50$ ms. Right: Time series of the 256 largest sources above the horizontal black line at source number 256. Below this line are 256 randomly selected sources of the left overs to indicate the noise level.

amount of noise present in the estimated solution. Note, that since this figure shows the simulated sources, only a minor part of the most active sources are active. The corresponding simulated EEG consists of the clean EEG constructed with a 3-spheres forward model by SPM5¹ with tissue conductivities brain:skull:scalp = 0.33:0.0041:0.33 S/m (ratio 1:1/80:1) and dimension 128×7204 added with 'realistic EEG noise'. The amount of noise added to the clean EEG signal is in the order of a SNR = 10 similar to [6], with SNR defined as the ratio between the power of the clean EEG and the noise. The noise is adopted from an evoked EEG study described in Sec. 3.2, where we have used the pre-stimulus period as noise. For source estimation three different 3-spheres head models with a lower spatial resolution than the 'true' forward model is used, and two of these are added with conductivity errors. In Fig. 3 both the location and the time series of the estimated source densities for MN, ARD, and SOFOMORE are shown. The estimated sources illustrated on the SPM glass brain correspond to the time point $t=50$ ms. In this figure a forward model with 'erroneous' conductivity value for the skull (0.0223 S/m) has been used. The MN estimate leads to a quite good localization of the main activity in the occipital lobe, however, also minor activity in the right pre-motor area is reconstructed. In the image with the time series included, it is seen that the MN solution also reconstructs quite a bit of activity outside the time period 25-75ms where no activity should be present. Compared to the MN solution, the source reconstruction using the ARD model is very sparse with a dominant source located correctly in left occipital region. However, the ARD model also captures minor activity in the pre-motor area. It is noted that the amplitude of the active sources are much higher than the simulated ones, since only a few account for most of the energy in the EEG signal. In contrast, the SOFOMORE model only reconstructs

¹The forward model was estimated by the SPM5 academic software (<http://www.fil.ion.ucl.ac.uk/spm/>), based on routines from BrainStorm (<http://neuroimage.usc.edu/brainstorm/>).

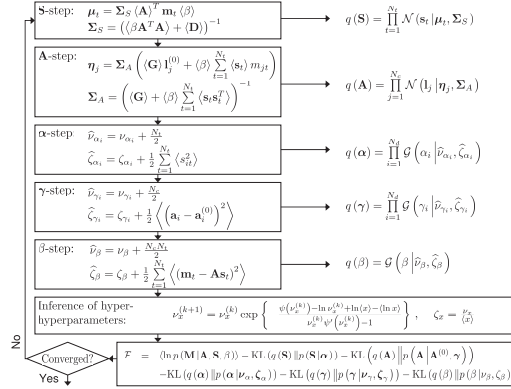


Fig. 1. Flow chart of the SOFOMORE model in a VB formulation. A fully factorized model in $\theta = \{\mathbf{S}, \mathbf{A}, \alpha, \gamma, \beta\}$ is applied, where each of the VB updates can be seen with their corresponding approximate posterior distribution to the right. The log evidence of the model is denoted \mathcal{F} , which includes Kullback-Leibler (KL) divergences of the approximate posterior distributions and their prior distributions. Following notations are used $\langle \cdot \rangle$ denoting the expectation, $\mathbf{I}_j^{(0)}$ the j 'th row in \mathbf{A} , and \mathbf{G} a diagonal matrix with λ in the diagonal. Furthermore, $\psi(\cdot)$ is the digamma function and its derivative is $\psi'(\cdot)$. Hyper-hyperparameters of the form ν_x and ζ_x are updated using Newton-Raphson.

the activity in the left occipital and minimal activity outside the time period of the simulated source is reconstructed. It is seen that regarding the forward fields as stochastic processes leads to an improvement relative to the source estimates obtained with fixed forward model in the ARD model.

In Fig. 4 similar source estimates with a 'correct' choice of conductivities are shown. A better estimate of the source amplitudes is shown compared to the simulated sources except for the ARD method due to its sparse nature. Indeed, an improvement of the ARD source estimates is also obtained by taking the uncertainty of the forward fields into account here, even with the 'correct' conductivity values given. This is further validated by the accuracy metrics described in Tab. 1. Table 2 shows the results for all three forward models used for source reconstruction. Note that illustrations as Fig. 3 and 4 for the third forward model with conductivity errors included (brain:skull:scalp=0.33:0.0013:0.33S/m) have been left out, since the results are quite similar to the ones in Fig. 3 and 4 with reconstruction of minor misleading activity in the pre-motor area for the MN and ARD methods. The differences can be seen from the validation metrics.

Table 1. Validation metrics.

Definition
$\text{MSE} = \sum_{t=1}^{N_t} \ \hat{\mathbf{s}}_t - \mathbf{s}_t\ ^2 / \sum_{t=1}^{N_t} \ \mathbf{s}_t\ ^2$
$\text{VE} = \sum_{j=1}^{N_s} \text{var}(\mathbf{M}_{j\cdot} - \mathbf{a}_j^T \mathbf{S}) / \sum_{j=1}^{N_s} \text{var}(\mathbf{M}_{j\cdot})$
$\text{DF} = \sum_{i \in \Theta} \ \hat{\mathbf{S}}_i - \mathbf{S}_i\ ^2 / \sum_{i \in \Theta} \ \mathbf{S}_i\ ^2$
AUC: Corresponding to $p(\hat{\mathbf{E}}(i') > \hat{\mathbf{E}}(i))$, with i' denoting the index of an active source and i a inactive source and $\hat{\mathbf{E}}(i) = \ \hat{\mathbf{S}}_i\ ^2 / \max(\ \hat{\mathbf{S}}\ ^2)$

3.2. Real EEG data

In this section real EEG data is used, which is from a multi-modal study on face perception, where faces and scrambled faces were presented for 600ms every 3600ms to a subject. A detailed description of the experiment is given in [15] and <http://www.fil.ion.ucl.ac.uk/spm> (where the data is available for download). In this contribution we reconstruct the average event related potential (ERP) of trials

Table 2. Results of validation metrics for MN, ARD, and SOFOMORE on three forward models with different conductivity ratios. SOFOMORE performs best in the metrics MSE, VE and AUC. Optimal value for VE is 88.2%. Note that MSE measures the ability to reconstruct the true active sources while keeping estimates of the true inactive sources small. SOFOMORE's ability to suppress the noise results in the best MSE. In contrast, DF measures only the relative squared error of the source estimates for the simulated region. Since the amplitudes of the MN estimates are higher than SOFOMORE's but still smaller than the true, MN perform best in this measure.

	MN			ARD			SOFOMORE		
	1:1/15:1	1:1/80:1	1:1/250:1	1:1/15:1	1:1/80:1	1:1/250:1	1:1/15:1	1:1/80:1	1:1/250:1
MSE	0.85	1.08	5.05	3.92	20.21	119.86	0.91	0.76	0.74
VE	99.0%	98.8%	98.8%	99.5%	99.1%	98.7%	95.5%	91.6%	81.5%
DF	0.76	0.42	0.30	2.11	8.03	78.44	0.88	0.69	0.56
AUC	0.93	0.97	0.98	0.89	0.93	0.87	0.92	0.99	0.99

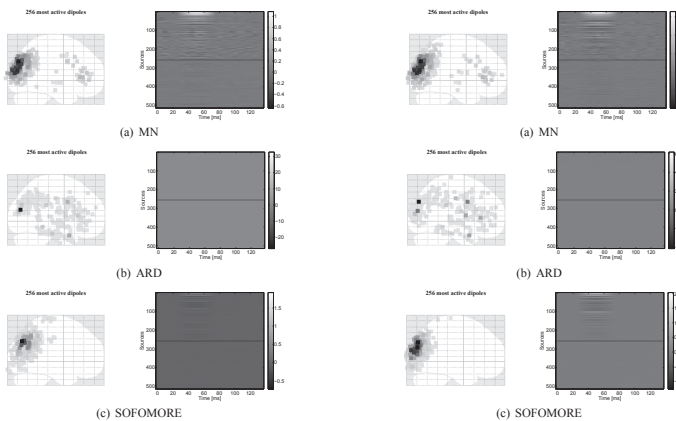


Fig. 3. Source estimates when forward model with 'erroneous' conductivity values (ratios 1:1/15:1) are used. Left: Activity at $t=50\text{ms}$. Right: Time series as indicated in Fig.2.

involving real faces as stimuli. The estimated source densities at $t=170\text{ms}$ for MN, ARD, and SOFOMORE are illustrated in Fig. 5. Quite different results are obtained, with the three source localization methods. The MN leads to activity in a larger area compared to the ARD and SOFOMORE, with most of its activity located in the right occipital region, the right frontal region and right fusiform gyrus. The ARD results in scattered activity with two prominent voxel in the left and right temporal lobe, respectively, and less prominent voxels in the visual cortex. In contrast, forward modeling in SOFOMORE leads to improved localization of activity in the visual cortex compared to the ARD with a

Fig. 4. Similar to Fig. 3 however with 'correct' conductivity values (ratios 1:1/80:1).

fixed forward model. The source estimate in SOFOMORE is quite similar to the MN solution, however with a dominant region in the left visual cortex and basically no frontal activity. The SOFOMORE also leads to weak activity in both the left and right fusiform gyrus, which is known to be connected with face processing. This is well-aligned with results reported in [10].

4. CONCLUSION

We presented the first results of a hierarchical Bayesian framework for simultaneous source and forward model reconstruction, with no explicit physical assumptions about the

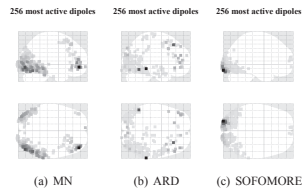


Fig. 5. Estimated activity $t = 170\text{ms}$ after stimulus. Tissue conductivities $\text{brain:skull:scalp} = 0.33:0.0041:0.33\text{S/m}$ are used.

source and forward model error. Both simulations and results from real EEG data illustrated an improvement of a similar model where the forward model is assumed fixed. Moreover, a comparison with the minimum norm method was also used to illustrate the applicability of the extended hierarchical model. Simulation results showed that the SOFOMORE model was able to reduce large distance errors. A serious concern of extending source localization methods to include forward model reconstruction is overfitting. However, the choice of an ARD prior on the forward fields allows corrections to mainly be performed where it is required to fit the signal while keeping the model simple.

5. REFERENCES

- [1] S. Baillet, J.C. Mosher, and R.M. Leahy, "Electromagnetic brain mapping," *IEEE Signal Processing Magazine*, vol. 18, pp. 14–30, 2001.
- [2] J. C. Mosher, R. M. Leahy, and P. S. Lewis, "EEG and MEG: forward solutions for inverse methods," *IEEE Transactions on Biomedical Engineering*, vol. 46, no. 3, pp. 245–259, 1999.
- [3] C.H. Wolters, A. Anwander, X. Tricoche, D. Weinstein, M.A. Koch, and R.S. MacLeod, "Influence of tissue conductivity anisotropy on EEG/MEG field and return current computation in a realistic head model: A simulation and visualization study using high-resolution finite element modeling," *Neuroimage*, vol. 30, no. 3, pp. 813–826, 2006.
- [4] S. Baillet and L. Garnero, "A bayesian approach to introducing anatomo-functional priors in the EEG/MEG inverse problem," *IEEE Transactions on Biomedical Engineering*, vol. 44, no. 5, pp. 374–385, 1997.
- [5] C. Phillips, M.D. Rugg, and K.J. Friston, "Anatomically Informed Basis Functions for EEG Source Localisation: Combining Functional and Anatomical Constraints," *NeuroImage*, vol. 16, no. 3, pp. 678–695, 2002.
- [6] M.-A. Sato, T. Yoshioka, S. Kajihara, K. Toyama, N. Goda, K. Doya, and M. Kawato, "Hierarchical bayesian estimation for MEG inverse problem," *Neuroimage*, vol. 23, no. 3, pp. 806–26, November 2004.
- [7] N.J. Trujillo-Barreto, E. Aubert-Vazquez, and W.D. Penny, "Bayesian M/EEG source reconstruction with spatio-temporal priors," *NeuroImage*, vol. 39, pp. 318–335, 2008.
- [8] J. C. Mosher, P. S. Lewis, and R. M. Leahy, "Multiple dipole modeling and localization from spatio-temporal MEG data," *IEEE Transactions on Biomedical Engineering*, vol. 39, no. 6, pp. 541–557, 1992.
- [9] M.S. Hämäläinen and R.J. Ilmoniemi, "Interpreting magnetic fields of the brain: minimum norm estimates," *Med. Biol. Eng. Comput.*, vol. 32, pp. 35–42, 1994.
- [10] K. Friston, L. Harrison, J. Daunizeau, S. Kiebel, C. Phillips, N. Trujillo-Barreto, R. Henson, G. Flandin, and J. Mattout, "Multiple sparse priors for the M/EEG inverse problem," *NeuroImage*, vol. 39, pp. 1104–1120, 2008.
- [11] C.M. Bishop, *Pattern Recognition and Machine Learning*, Springer, NY 10013 (USA), 2006.
- [12] S.M. Plis, J.S. George, S.C. Jun, D.M. Ranken, P.L. Volegov, and D.M. Schmidt, "Probabilistic forward model for electroencephalography source analysis," *Physics in Medicine and Biology*, vol. 52, no. 17, pp. 5309–5328, 2007.
- [13] S. Lew, C. Wolters, A. Anwander, S. Makeig, and R.S. MacLeod, "Low resolution conductivity estimation to improve source localization," in *New Frontiers in Biomagnetism. Proc. of the 15th Int. Conf. on Biomag.*, 2007, vol. 1300 of *Int. Congress Series*, pp. 149–152.
- [14] C. Stahlhut, M. Mørup, O. Winther, and L.K. Hansen, "SOFOMORE: Combined EEG source and forward model reconstruction," in *6th IEEE Int. Symp. on Biomed. Imaging (ISBI): From Nano to Macro*, 2009, pp. 450–453.
- [15] R.N.A. Henson, Y. Goshen-Gottstein, T. Ganel, L.J. Otten, A. Quayle, and M.D. Rugg, "Electrophysiological and hemodynamic correlates of face perception, recognition and priming," *Cerebral Cortex*, vol. 13, pp. 793–805, 2003.

APPENDIX C

Simultaneous EEG Source and Forward Model Reconstruction (SOFOMORE) using a Hierarchical Bayesian Approach

Carsten Stahlhut, Morten Mørup, Ole Winther, and Lars K. Hansen. Simultaneous EEG Source and Forward Model Reconstruction (SOFOMORE) using a Hierarchical Bayesian Approach. *Journal of Signal Processing System (JSPS)*. October 2010.

J Sign Process Syst
DOI 10.1007/s11265-010-0527-0

Simultaneous EEG Source and Forward Model Reconstruction (SOFOMORE) Using a Hierarchical Bayesian Approach

Carsten Stahlhut · Morten Mørup ·
Ole Winther · Lars Kai Hansen

Received: 15 January 2010 / Revised: 23 April 2010 / Accepted: 3 September 2010
© Springer Science+Business Media, LLC 2010

Abstract We present an approach to handle forward model uncertainty for EEG source reconstruction. A stochastic forward model representation is motivated by the many random contributions to the path from sources to measurements including the tissue conductivity distribution, the geometry of the cortical surface, and electrode positions. We first present a hierarchical Bayesian framework for EEG source localization that jointly performs source and forward model reconstruction (SOFOMORE). Secondly, we evaluate the SOFOMORE approach by comparison with source reconstruction methods that use fixed forward models. Analysis of simulated and real EEG data provide evidence that reconstruction of the forward model leads to improved source estimates.

Keywords EEG · Inverse problem · Source localization · Distributed models · Variational Bayes · Forward model reconstruction

Thanks to the Lundbeck Foundation for supporting this work through the Center for Integrated Molecular Brain Imaging (CIMBI)—www.cimbi.org.

C. Stahlhut (✉) · M. Mørup · O. Winther · L. K. Hansen
Department of Informatics and Mathematical Modelling,
Technical University of Denmark, Richard Petersens Plads,
2800 Kgs. Lyngby, Denmark
e-mail: cs@imm.dtu.dk

M. Mørup
e-mail: mm@imm.dtu.dk

O. Winther
e-mail: owi@imm.dtu.dk

L. K. Hansen
e-mail: lkh@imm.dtu.dk

1 Introduction

Electroencephalography (EEG) measures the electrical fields that arise from neural activity in the brain by an array of sensors placed on the scalp. We are interested in the reconstruction of the EEG source generators, which is known as an severe ill-posed inverse problem with a mapping from source space to sensor space that is many to one. Despite the ill-posed nature of EEG source reconstruction it represents a key problem of high current interest for the functional imaging community due to the EEG signal's high temporal resolution and relatively direct coupling to the neural signal, compared to, e.g., functional magnetic resonance imaging (fMRI) and positron emission tomography (PET). Under the quasi-static approximation to Maxwell's equations the relation between the measured EEG signal and the brain's current sources can be expressed as a linear instantaneous form in the sources. In particular, if the measured EEG signal is denoted $\mathbf{M} \in \mathbb{R}^{N_c \times N_t}$ and the current sources $\mathbf{S} \in \mathbb{R}^{N_d \times N_t}$, and with N_c , N_d , and N_t being the number of channels, dipoles, and time samples, respectively, the forward relation can be written [3],

$$\mathbf{M} = \mathbf{A}\mathbf{S} + \mathcal{E}, \quad (1)$$

where we further have assumed the noise \mathcal{E} additive. The interrelationship between the sensors and the current sources is given by the lead field matrix/forward model $\mathbf{A} \in \mathbb{R}^{N_c \times N_d}$ with the rows referred to as the lead fields for the sensors and the columns as the forward fields for the sources. In this contribution we will, for simplicity, assume local dipoles to be perpendicular to cortex. However, a more flexible orientation can easily

be incorporated by inclusion of three columns in \mathbf{A} per dipole corresponding to the three directional components. Different levels of complexity of the head model exist, where the spherical head model, the boundary element method (BEM), and finite element methods (FEM) are examples of models of increasing complexity [15, 26].

To regularize the inverse EEG problem additional information based in anatomy and physiology and formulated as statistical priors or mathematical constraints are needed [2, 18, 20, 24]. This has led to the development of a large number of inverse methods in the literature. One approach restricts the estimated current sources to a limited number of current dipoles. This can lead to an over-determined inverse problem where the limited set of dipoles are fitted to the data by minimizing a measure of reconstruction error such that the location, orientation, and amplitude of the dipoles can be obtained [16, 21]. A different class of inverse methods is the so-called distributed models also known as source imaging techniques, where the spatial distribution of the current sources is recovered. The well-known minimum norm (MN) estimate [11] reconstructs a source distribution that minimizes the l_2 -norm of the measurements and the explained EEG by a given source configuration.

Relevant ‘A priori’ information on the sources \mathbf{S} can conveniently be injected within a Bayesian framework through a prior distribution $p(\mathbf{S})$. In fact, it can be argued that it is the very choice of source prior that differentiates current source localization methods. Given the observed data we can now use the marginal posterior distribution as representation of the unknown sources,

$$p(\mathbf{S}|\mathbf{M}) = \frac{p(\mathbf{M}, \mathbf{S})}{p(\mathbf{M})} = \frac{\int p(\mathbf{M}|\boldsymbol{\theta}) p(\boldsymbol{\theta}) d\boldsymbol{\theta}_{\mathbf{S}}}{p(\mathbf{M})} \quad (2)$$

where $\boldsymbol{\theta}$ is a set of parameters for the model including \mathbf{S} . We use the notation $\boldsymbol{\theta}_{\mathbf{S}}$ for the parameter set $\boldsymbol{\theta}$ excluding the sources \mathbf{S} . To obtain the posterior distribution for \mathbf{S} we marginalize all other parameters than \mathbf{S} . The distribution $p(\mathbf{M})$ is the marginal likelihood also referred to as the model evidence and is given by $p(\mathbf{M}) = \int p(\mathbf{M}|\boldsymbol{\theta}) p(\boldsymbol{\theta}) d\boldsymbol{\theta}$. The model evidence is important for model comparison. As parameters are integrated out in the model evidence, the Bayesian framework naturally favors simpler models over more complex models which A priori are able to model a larger range of data sets. Unfortunately, the marginal likelihood is intractable for most prior distributions, and thereby the posterior $p(\mathbf{S}|\mathbf{M})$ cannot be computed

either. Thus, approximate methods, e.g., the Laplace approximation [7], variational Bayes (VB) [4, 5, 20], and Markov chain Monte Carlo (MCMC) sampling [19] have been used.

In contrast to existing source localization methods we will avoid to make the assumption that the forward model is known and fixed. This is motivated by the many noise processes that contribute to the forward model, including the geometrical representation of the cortical surface, the conductivity distribution, and electrode positions. When ‘realistic head models’ are constructed from tissue segmentation based on e.g., structural MRI, the geometry of the head model is affected by the resolution and tissue segmentation errors. These errors can be presented as small magnitude perturbations of the head model shape [25].

The issue of modeling the forward process was recently pursued in [13, 19] from quite a different viewpoint than ours. In [13] the basic structure of the forward model is assumed known, while the skull-brain conductivity ratio is an unknown parameter which is estimated simultaneously with a few sources in a dipole fit setting. Similarly, [19] has proposed a probabilistic approach for distributed models to account for uncertainties in the skull conductivity. Here, we apply a more unconstrained approach by modeling the whole forward model as uncertain using a prior distribution for the forward model. We propose a first attempt for distributed models to perform simultaneous source and forward model reconstruction, in short the SOFOMORE model [22, 23]. As an alternative to the SOFOMORE model the total least squares (TLS) [9] can be mentioned, which is an extension of the ordinary least squares problem such that it also can deal with errors in \mathbf{A} .

An illustration of the concept of applying a stochastic forward model in SOFOMORE is given in Fig. 1, in

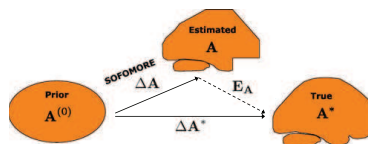


Figure 1 Illustration of the concept of forward model reconstruction. The SOFOMORE model uses a potential erroneous forward model $\mathbf{A}^{(0)}$ as prior mean and depending on the regions of source activity the model can flexibly correct these regions by $\Delta\mathbf{A}$. As such, the estimated forward model is \mathbf{A} with error \mathbf{E}_A to the true underlying forward model denoted \mathbf{A}^* .

which the conventional forward model $\mathbf{A}^{(0)}$ is represented as 'the prior mean'. Due to the uncertainties mentioned above, we expect a difference between the true underlying model of the forward fields, \mathbf{A}^* and $\mathbf{A}^{(0)}$. The main novelty of our SOFOMORE approach is thus the posterior distribution of the forward model \mathbf{A} representing our uncertainty of the deviations of the prior mean $\mathbf{A}^{(0)}$.

In this paper we focus on the conceptual development and the statistical aspects of the SOFOMORE model. Furthermore, we evaluate the performance of the stochastic forward propagation model and compare it to the MN method at different signal-to-noise ratios (SNR). The proposed model may be considered as a first step for modeling uncertainty in \mathbf{A} by assuming the errors to be Gaussian distributed. The next step may be a more realistic model of uncertainty arising from sensor placement and head shape. In such a model one would let the uncertainty in an element of \mathbf{A} depend non-linearly upon the distance from the sensor to the source [15]. Part of this work has appeared in [22].

2 Methods

Given the linear relation in Eq. 1 and assuming the noise to be time independent Gaussian distributed, the observation model can be written as, $p(\mathbf{m}_t | \mathbf{s}_t, \Sigma_{\mathcal{E}}) = \mathcal{N}(\mathbf{m}_t | \mathbf{A}\mathbf{s}_t, \Sigma_{\mathcal{E}})$ where $\Sigma_{\mathcal{E}}$ is the noise spatial covariance matrix. Temporal correlations could be incorporated as outlined in [7]. In the remainder of this paper we also assume spatial independence $\Sigma_{\mathcal{E}} = \beta^{-1} \mathbf{I}_{N_s}$.

2.1 The Minimum Norm Estimate

In a minimum norm setting a multivariate Gaussian prior for the sources with zero mean and covariance $\alpha^{-1} \mathbf{I}_{N_s}$ is assumed. Moreover, it is assumed that the forward propagation model $\mathbf{A} = \mathbf{A}^{(0)}$ is known. With Bayes rule as in Eq. 2 it is seen that the posterior distribution is maximized by

$$\Sigma_{\mathbf{M}} = (\alpha^{-1} \mathbf{A}^{(0)} \mathbf{A}^{(0)T} + \beta^{-1} \mathbf{I}_{N_s})^{-1} \quad (3)$$

$$\mathbf{S}_{\text{MN}} = \alpha^{-1} \mathbf{A}^{(0)T} \Sigma_{\mathbf{M}} \mathbf{M}. \quad (4)$$

It is noted that since the likelihood and prior are both Gaussian distributions, the posterior and marginal likelihood will also be Gaussian distributions. The estimation of the sources, the precision parameters α and β are performed using a standard expectation-maximization (EM) scheme [5].

2.2 SOFOMORE: Hierarchical Model

We propose a hierarchical model for the forward fields and EEG sources,

$$\begin{aligned} \mathbf{m}_t &\sim \mathcal{N}(\mathbf{m}_t | \mathbf{A}\mathbf{s}_t, \beta^{-1} \mathbf{I}_{N_s}) \\ \mathbf{s}_t &\sim \mathcal{N}(\mathbf{s}_t | \mathbf{0}, \mathbf{D}^{-1}), \quad \mathbf{D} = \text{diag}(\boldsymbol{\alpha}) \\ \mathbf{a}_i &\sim \mathcal{N}(\mathbf{a}_i | \mathbf{a}_i^{(0)}, \gamma_i^{-1} \mathbf{I}_{N_s}) \\ \alpha_i &\sim \mathcal{G}(\alpha_i | v_{\alpha_i}, \zeta_{\alpha_i}) \\ \gamma_i &\sim \mathcal{G}(\gamma_i | v_{\gamma_i}, \zeta_{\gamma_i}) \\ \beta &\sim \mathcal{G}(\beta | v_{\beta}, \zeta_{\beta}) \end{aligned} \quad (5)$$

For the current sources we assign a zero-mean multivariate Gaussian prior distribution with a diagonal precision matrix \mathbf{D} , where the diagonal elements are $\boldsymbol{\alpha} = \{\alpha_i\}_{i=1}^{N_s}$, thus, the model allows dipoles to have different variance. The current variances are assumed not to change over the time period $t = 1 : N_t$ and are modeled by an automatic relevance determination-type prior (ARD) quite similar to [20], i.e., using a conjugate prior for the precision parameters $\boldsymbol{\alpha}$. The conjugate prior for the precision parameter is the Gamma distribution denoted $\mathcal{G}(\cdot)$ with shape parameter v_{α_i} and rate parameter ζ_{α_i} such that the mean is given by $v_{\alpha_i} / \zeta_{\alpha_i}$ and the variance $v_{\alpha_i} / \zeta_{\alpha_i}^2$. Assigning an ARD prior to the sources is a key aspect in preventing overfitting, since by optimization of the hyperparameters it is possible to automatically drive a proportion of the sources towards zero, see e.g., [5] for a detailed explanation of the ARD property.

As we regard the forward model as a stochastic process, the conventional forward propagation model $\mathbf{A}^{(0)}$ is used as a mean value in a multivariate Gaussian prior. Each of the forward fields are modeled as independent, such that the prior distribution of a single forward field is given by prior mean $\mathbf{a}_i^{(0)}$ (ith column in $\mathbf{A}^{(0)}$) and precision γ_i , i.e., we have $\boldsymbol{\gamma} = \{\gamma_i\}_{i=1}^{N_s}$. Assigning each of the forward fields a separate precision parameter allows us to differentiate the individual amounts of correction to forward fields. The precision parameters for the forward fields are also modeled by conjugate prior distributions, i.e., Gamma distributions with shape and rate parameter, v_{γ_i} and ζ_{γ_i} , respectively. Likewise, the inverse noise variance is assumed to be Gamma distributed, $\mathcal{G}(\beta | v_{\beta}, \zeta_{\beta})$. We note that the MN estimate is a special case of the SOFOMORE solution, where the prior for the forward fields are just delta functions centered at $\mathbf{a}_i^{(0)}$ and, moreover, the values of precisions for the current sources are all equal.

Given the hierarchical structure of the SOFOMORE model with the parameters $\theta = \{\mathbf{S}, \mathbf{A}, \alpha, \gamma, \beta\}$ the marginal posterior distribution of the current sources becomes analytically intractable, thus, approximations are needed. Here, we apply a standard VB framework [1, 5, 8], in which a parameterized simpler distribution $q(\theta)$ approximates the true joint posterior $p(\theta | \mathbf{M})$. In the VB framework the parameters of the $q(\theta)$ distribution are determined with a maximization of a lower bound of the marginal likelihood obtained by the Jensen's inequality,

$$\mathcal{L} = \ln \int p(\mathbf{M} | \theta) p(\theta) d\theta \geq \int q(\theta) \ln \frac{p(\mathbf{M} | \theta) p(\theta)}{q(\theta)} d\theta = \mathcal{F}_{\mathcal{M}}(q(\theta)). \quad (6)$$

$\mathcal{F}_{\mathcal{M}}$ is the lower bound for model \mathcal{M} and θ is the union of the parameter sets $\theta = \{\mathbf{S}, \mathbf{A}, \alpha, \gamma, \beta\}$ such that the joint distribution $p(\mathbf{M}, \theta)$ is given by

$$p(\mathbf{M}, \theta) = p(\mathbf{M} | \mathbf{S}, \mathbf{A}, \beta) p(\mathbf{S} | \alpha) p(\mathbf{A} | \gamma) p(\alpha | \nu_\alpha, \zeta_\alpha) \cdot p(\beta | \nu_\beta, \zeta_\beta) p(\gamma | \nu_\gamma, \zeta_\gamma) \quad (7)$$

In order to maximize the lower bound of the log marginal likelihood we further assume factorization in the parameters $\theta = \{\mathbf{S}, \mathbf{A}, \alpha, \gamma, \beta\}$ of the variational posterior $q(\theta)$, i.e., the full variational posterior distribution $q(\theta)$ is restricted to

$$q(\theta) = q(\mathbf{S}) q(\mathbf{A}) q(\alpha) q(\beta) q(\gamma). \quad (8)$$

The factorized variational distribution is a crude approximation and it is mainly motivated by computational convenience as it provides for a tractable marginal posterior distribution. In turn each of the variational marginal posterior distributions are maximized by solving $\delta \mathcal{F}_{\mathcal{M}} / \delta q(\mathbf{Z}) = 0$ to obtain [1],

$$q(\mathbf{Z}) \propto e^{(\ln p(\mathbf{M}, \theta))_{q(\mathbf{Z})}}, \quad (9)$$

where $q(\mathbf{Z})$ denotes one of the variational marginal posterior distributions in Eq. 8 and the expectation $\langle \cdot \rangle$ is taken w.r.t. $q(\theta_{\setminus \mathbf{Z}})$ with $\theta_{\setminus \mathbf{Z}}$ denoting the set θ except \mathbf{Z} . Following this scheme we obtain the sequential VB updates as illustrated in Fig. 2, in which each of the variational marginal posterior distributions of the

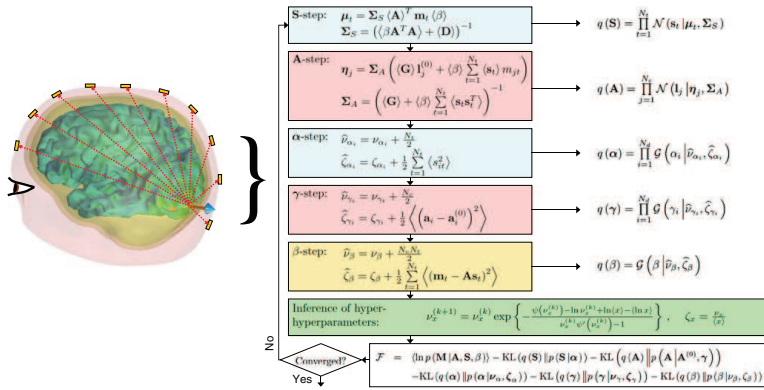


Figure 2 Flow chart of the SOFOMORE model in a VB formulation. A fully factorized model in $\theta = \{\mathbf{S}, \mathbf{A}, \alpha, \gamma, \beta\}$ is applied, where each of the VB updates can be seen with their corresponding approximate posterior distribution to the right. The log evidence of the model is denoted \mathcal{F} , which include Kullback–Leibler (KL) divergences of the approximate posterior distributions and their prior distributions. Following notations are used $\langle \cdot \rangle$ denoting the expectation, $\mathbf{I}_J^{(0)}$ the j th row in \mathbf{A} , and \mathbf{G} a diagonal matrix with λ in the diagonal. Furthermore, $\psi(\cdot)$ is the digamma

function and its derivative is $\psi'(\cdot)$. Hyper-hyperparameters of the form ν_x and ζ_x are updated using Newton–Raphson. The colorcoding light blue, red, and yellow correspond to the blue, red, and yellow colorcoding to the left, i.e., the source, forward fields, and channels, respectively. The green colorcoding of the step with inference of hyper-hyperparameters means that both hyper-hyperparameters associated with sources, forward fields, and channels are updated here.

parameters are given. A more detailed derivation of the VB updates for the SOFOMORE model is provided in Appendix A.

3 Experiments

We demonstrate the viability of the SOFOMORE model source reconstruction for both simulated and real EEG data. We present the recovered source density estimates both with and without estimated forward fields at different noise levels. Besides the MN method with a fixed forward model, we also compare the SOFOMORE model to a model with a similar hierarchical structure as the SOFOMORE, however, with a fixed forward field, i.e., no \mathbf{A} - or γ -steps are performed. We denote this model as the ARD model, due to its ARD prior on the sources. For clarity of the results, the methods work directly on the measurements \mathbf{M} , i.e., no pre-processing is performed. As validation metrics we use the mean square error (MSE), variance explained (VE), degree of focalization (DF), and area under the receiver operating curves (AUC). Definition of the validation metrics is given in Table 1 in Section 3.1.

3.1 Simulations

In our simulations a small cortical area in the left occipital lobe is simulated as active. The source signal consists of a half sine of duration 50 ms starting at $t = 25$ ms quite similar to [6, 10, 14]. The simulated sources at $t = 50$ ms are shown in the SPM glass-brain representation in Fig. 3. Due to the mapping from cortex to the glass-brain representation, minor activity seems to appear at the inner part of left hemisphere. Moreover, Fig. 3 shows the time series of the simulated sources. The first 256 rows in the image are the 256 most active sources and the last 256 rows below the black horizontal line in the image are 256 randomly drawn sources with the 256 most active sources excluded. This will give an idea of the amount of noise present in the estimated solution.

The corresponding simulated EEG consists of the clean EEG constructed with a three-spheres forward model by SPM5¹ with tissue conductivities brain:skull:scalp = 0.33:0.0041:0.33S/m (ratio 1:1/80:1) and dimension $128 \times 7,204$ added with 'realistic EEG noise'. In order to illustrate the robustness of the SOFOMORE

Table 1 Validation metrics with MSE measuring the ability to reconstruct the amplitudes of the true sources, VE the amount of variance explained in the observed data \mathbf{M} , i.e., signal and noise.

Definition	
MSE	$\sum_{i=1}^{N_t} \ \hat{\mathbf{s}}_i - \mathbf{s}_i\ ^2 / \sum_{i=1}^{N_t} \ \mathbf{s}_i\ ^2$
VE	$1 - \sum_{j=1}^{N_s} \text{var}(\mathbf{M}_j - \mathbf{I}_j^T \hat{\mathbf{S}}) / \sum_{j=1}^{N_s} \text{var}(\mathbf{M}_j)$
DF	$\sum_{i \in \Theta} \ \hat{\mathbf{s}}_i - \mathbf{s}_i\ ^2 / \sum_{i \in \Theta} \ \mathbf{s}_i\ ^2$
AUC	Corresponding to $p(\hat{\mathbf{E}}(i') > \hat{\mathbf{E}}(i))$, with i' denoting the index of an active source and i a inactive source and $\hat{\mathbf{E}}(i) = \ \hat{\mathbf{s}}_i\ ^2 / \max(\ \hat{\mathbf{s}}\ ^2)$

In contrast to MSE, DF measures only the relative squared error of the source estimates for the simulated region Θ . AUC measures the area under the ROC curve.

model, the amount of noise added to the clean EEG signal is varied. More specifically we evaluate the source reconstruction at the SNR = {0.5, 1, 3, 5, 8, 10}, with

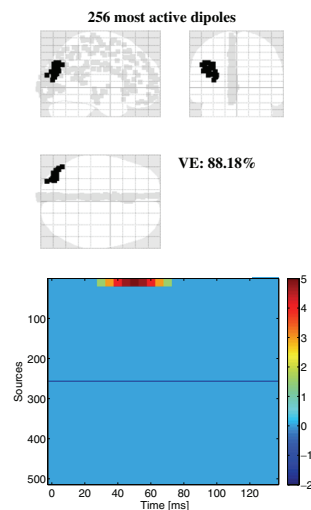


Figure 3 Top Simulated sources at $t = 50$ ms. Bottom Time series of the 256 largest sources above the horizontal black line at source number 256. Below this line 256 randomly selected sources of the left overs are seen which indicate the noise level. Given the simulated sources the amount of variance that should be explained in the observations, \mathbf{M} , is $VE = 88.18\%$.

¹The forward model was estimated by the SPM5 academic software (<http://www.fil.ion.ucl.ac.uk/spm/>), based on routines from BrainStorm (<http://neuroimage.usc.edu/brainstorm/>).

SNR defined as the ratio between the power of the clean EEG and the noise. The noise is adopted from an evoked EEG study described in Section 3.2, where the pre-stimulus period has been used as noise. For source estimation three different three-spheres head models with a lower spatial resolution than the ‘true’ forward model are used, and two of these are added with conductivity errors. We will denote these forward models as **A1**, **A2**, and **A3**, with conductivity values (brain:skull:scalp) specified as

- **A1**: 0.33:0.0223:0.33 S/m (ratio 1:1/15:1)
- **A2**: 0.33:0.0041:0.33 S/m (ratio 1:1/80:1)
- **A3**: 0.33:0.0013:0.33 S/m (ratio 1:1/250:1)

In Fig. 4 both the location and the time series of the estimated source densities for MN, ARD, and SOFOMORE are shown. The estimated sources illustrated on the SPM glass-brain correspond to the time point $t = 50$ ms. In this figure forward model **A1** with ‘erroneous’ conductivity value for the skull (0.0223 S/m) has been used.

The MN estimate leads to a quite good localization of the main activity in the occipital lobe, however, in addition we find spurious activity in the right pre-motor area. In the image with the time series included, it is seen that the MN solution also reconstructs noticeable activity outside the activation time period (25–75 ms). Compared to the MN solution, the source reconstruction using the ARD model is very sparse with a dominant source located correctly in left occipital region. However, the ARD model also finds activity in the pre-motor area. It is noted that the amplitude of the active sources are much higher than the simulated ones, since only a few account for most of the energy in the EEG signal. In contrast, the SOFOMORE model only reconstructs the activity in the left occipital and minimal activity outside the activation time period. Thus, the results provide evidence that reconstruction of the forward fields leads to improved source estimation relative to the conventional fixed forward model, as hypothesized. In Fig. 5 similar source estimates with a ‘correct’ choice of conductivities are shown. A better estimate

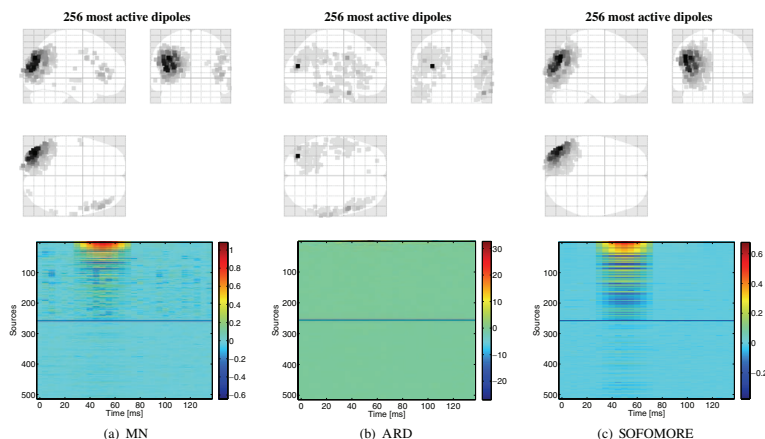


Figure 4 Source estimates when forward model **A1** with ‘erroneous’ conductivity values (ratios 1:1/15:1) is used and when the SNR = 10. *Upper row* Activity at $t = 50$ ms. *Lower row* Time series as indicated in Fig. 3. The MN reconstructs activity in the left occipital lobe correctly, however minor misleading activity is reconstructed in the right pre-motor area. The ARD leads to a very sparse solution with primarily one domination dipole

located correctly. Similarly, the ARD recover misleading activity in the pre-motor area and, moreover, the temporal lobe. In contrast the reconstructed activity by the SOFOMORE is located correctly in the left occipital lobe while simultaneously suppressing the noise, which can be seen by the glass brain maps and the time series map. Indeed the ARD source estimates are improved by applying a stochastic forward model in the SOFOMORE.

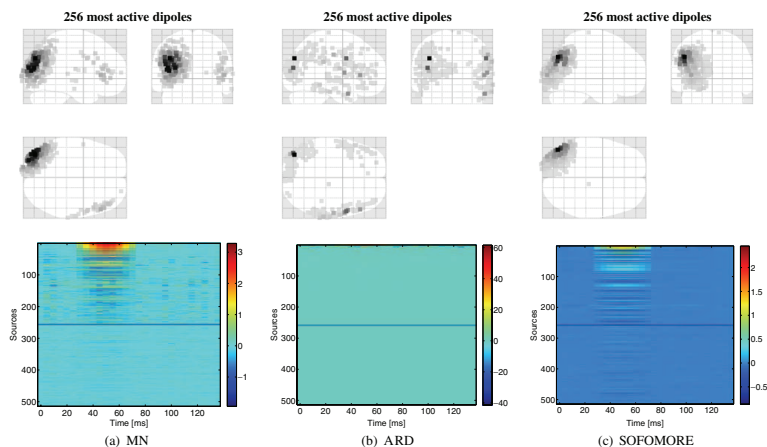


Figure 5 Source estimates when forward model **A2** with 'correct' conductivity values (ratios 1:1/80:1) is used and when the SNR = 10. Similar to Fig. 4 the *upper row* shows the activity at $t = 50$ ms and the *lower row* the time series as indicated in Fig. 3. Again the MN reconstructs activity correctly in the left occipital lobe but now with source amplitudes closer to the simulated ones and a little improvement in the suppression of noise. However, also

minor misleading activity is reconstructed in the right pre-motor area. With these conductivity values the ARD also leads to a very sparse solution with two domination dipoles located correctly. Once more the ARD recovers misleading activity in the pre-motor area and the temporal lobe. The SOFOMORE leads to source amplitudes closer to the simulated ones compared to the ones estimated in Fig. 4.

of the source amplitudes is seen compared to the simulated sources except for the ARD method due to its sparse nature. Indeed, an improvement of the ARD source estimates is also obtained by taking the uncertainty of the forward fields into account here, even with the 'correct' conductivity values given. This is further validated by the accuracy metrics described in Table 1. Figure 6 shows similar results as reported in Figs. 4 and 5, however now using the third forward model **A3** with conductivity errors included (brain:skull:scalp = 0.33:0.0013:0.33 S/m). Again the MN and ARD methods lead to minor misleading activity reconstructed in the pre-motor area.

Figure 7 shows the results for all three forward models used for source reconstruction. In Fig. 7a we have the MSE as a function of SNR. This metric measures the ability to reconstruct the true active current sources while keeping estimates of the true inactive sources small. In this figure it is seen that the SOFOMORE model in general for the three forward models per-

forms best in the MSE measure. However, the MN with the 'erroneous' conductivity values is actually also one of the best candidates for this measure. In contrast, DF only measures the relative squared error of the source estimates for the simulated region. Since the amplitudes of the MN estimates are higher than the SOFOMORE's yet smaller than the true, MN performs best in this measure, see Fig. 7b. The fact that the estimated sources for the MN overall have higher amplitudes is also reflected in the VE metric. Figure 7c reveals that the MN as a consequence describes almost the full simulated EEG, signal as well as noise. However, this type of overfitting is not as pronounced for the SOFOMORE, which is closer to the optimal VE values. The optimal VE value at the different SNR is represented by the solid black line, corresponding to the clean EEG signal. Finally, we have the AUC measure in Fig 7d, where the MN and SOFOMORE overall perform equally well for high SNR. However, for very small SNRs MN seems to be more robust in the AUC

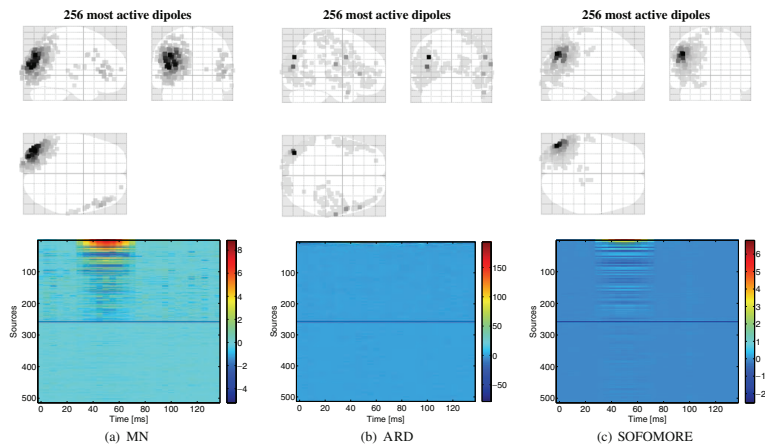


Figure 6 Source estimates when forward model **A3** with ‘erroneous’ conductivity values (ratios 1:1/250:1) is used and when the SNR = 10. *Upper row* Activity at $t = 50$ ms. *Lower row* Time series as indicated in Fig. 3. The MN reconstructs activity in the left occipital lobe correctly, however minor misleading activity

is reconstructed in the right pre-motor area. Once more the ARD leads to a very sparse solution with two domination dipoles located correctly. The ARD recovers misleading activity in the pre-motor area and the temporal lobe.

measure compared to the SOFOMORE. This may be due to the fact that SOFOMORE more or less can ‘turn off’ sources completely, and leave other spurious sources explain the signal and noise. Pre-processing can improve SNR and may be crucial at very low SNRs in order for the SOFOMORE to operate efficient w.r.t. to the AUC measure.

When comparing the validation metrics for the SOFOMORE with the ARD model we see that in all measures the SOFOMORE performs better than the ARD. This is a consequence of the ARDs sparse solutions, in particular for the metrics MSE, DF, VE, and AUC and at high SNRs.

3.1.1 Evaluation of the Change in Forward Model

In order to validate that the SOFOMORE captures the main errors in the forward model we analyze the difference between forward models **A1** and **A2**. Importantly, it should be recalled that **A2** is not the true forward model, but a low resolution resampling of the true forward model **A*** and with the same conductivity

values. However, this will give an indication of which brain areas that differs substantially. Figure 8 illustrates the squared difference between the forward fields in **A1** and **A2** relative to the squared length of the forward fields in **A2**, i.e., we have $\|\mathbf{a}_i^{(2)} - \mathbf{a}_i^{(1)}\|^2 / \|\mathbf{a}_i^{(2)}\|^2$ with $\mathbf{a}_i^{(1)}$ and $\mathbf{a}_i^{(2)}$ corresponding to the i th forward field in **A1** and **A2**, respectively. A very dominant left frontal region indicate a large mismatch between the two forward models. When inspecting the estimated parameter γ in the SOFOMORE it is indeed the left frontal region that the SOFOMORE model predicts as being the region with the largest uncertainty. This is shown in Fig. 8b, where the inverse precision values are illustrated. In Fig. 8c we analyze the change in the forward model performed in relation to the prior mean in this case **A1**, i.e., we have the squared difference between the estimated forward fields **A** and prior mean **A1** normalized with the squared length of the forward fields in **A2**. Note, that the change in the forward fields is mainly performed in the region where the activity arises from. Furthermore, a minor correction of the frontal forward field that was actually most affected is also performed.

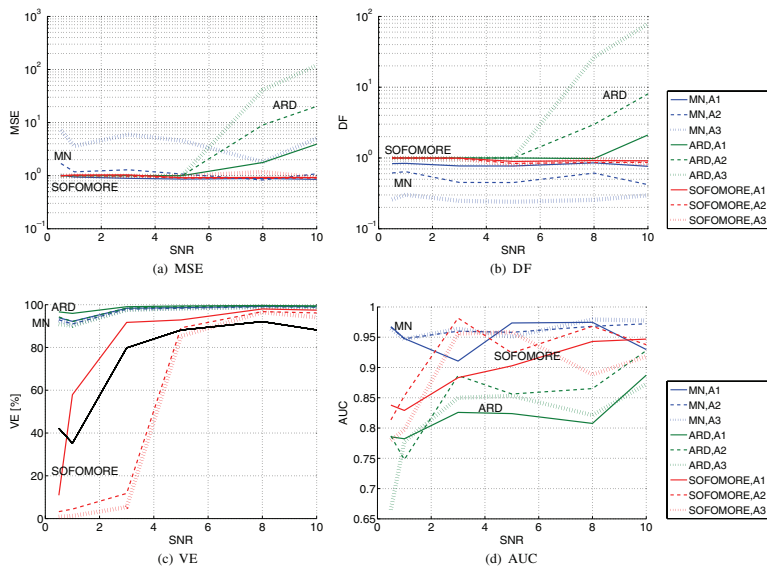


Figure 7 Results of validation metrics for MN, ARD, and SOFOMORE at different SNRs and on three forward models with different conductivity ratios. It is seen that the SOFOMORE model in general for the three forward models performs best in the MSE measure. However, the MN with the 'erroneous' conductivity ratios 1:1/15:1 (A1) is actually also one of the best candidates for this measure. Recall, that the MSE measures how

well the true active sources are reconstructed while keeping estimates of the true inactive sources small. This is indeed the case for the estimated sources by the SOFOMORE for all three forward models. In contrast, DF measures only the relative squared error of the source estimates for the simulated region. Since the amplitudes of the MN estimates are higher than SOFOMORE's but still smaller than the true, MN performs best in this measure.

3.2 Real EEG Data

In this section real EEG data from a multi-modal study on face perception is used for illustration. In the multi-modal study faces and scrambled faces were presented to the subject for 600 ms every 3,600 ms. A detailed description of the experiment is given in [12] and <http://www.fil.ion.ucl.ac.uk/spm> (where the data is also available for download). In this contribution we reconstruct the average event related potential (ERP) of trials involving real faces as stimuli. The estimated source densities at $t = 170$ ms for MN, ARD, and SOFOMORE are illustrated in Fig. 9. Quite different results are obtained with these three source localization

methods. The MN leads to activity in a larger area compared to the ARD and SOFOMORE, with most of its activity located in the right occipital region, the right frontal region and right fusiform gyrus. The ARD results in scattered activity with two prominent voxels in the left and right temporal lobe, respectively, and less prominent voxels in the visual cortex. In contrast, forward modeling in SOFOMORE leads to improved localization of activity in the visual cortex compared to the ARD with a fixed forward model. The source estimate in SOFOMORE is quite similar to the MN solution, however with a dominant region in the left visual cortex and basically no frontal activity. The SOFOMORE also leads to weak activity in both the

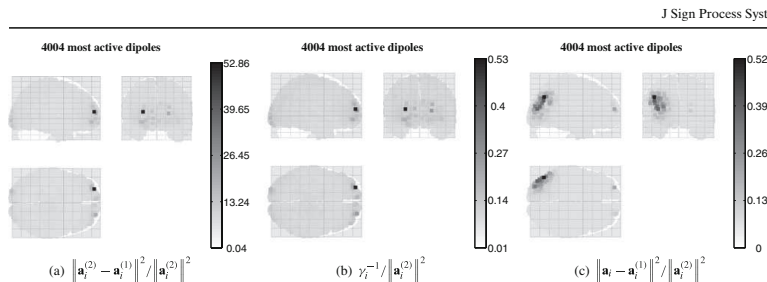


Figure 8 Comparison of simulated differences between forward models and the estimated differences. All subfigures are normalized with the squared norm of $\mathbf{A}\mathbf{2}$ to give a relative change of the forward fields. **a** This subfigure illustrates the relative squared difference in forward fields, $\Delta\mathbf{A}^* = \mathbf{A}\mathbf{2} - \mathbf{A}\mathbf{1}$. **b** Inverse values of estimated precision parameters γ in the SOFOMORE model with prior mean $\mathbf{A}\mathbf{1}$. It is seen that the dominating error

in forward fields in \mathbf{a} is captured by the precision parameter γ in the sense that it specifies that the given forward field is very uncertain (large variance). **c** The squared change in the forward fields $\Delta\mathbf{A} = \mathbf{A} - \mathbf{A}\mathbf{1}$. It is seen that the change in the forward fields is primarily performed for the region where the activity is located and, moreover, a minor change of the forward field with the largest error in \mathbf{a} .

left and right fusiform gyrus, which is known to be connected with face processing. These activated regions are well-aligned with the results reported in [7].

4 Conclusion

We have presented the first results of a hierarchical Bayesian framework for simultaneous detailed source and forward model reconstruction. Both in simulation and in analysis of a visual stimulus EEG data

set we found improvements in the source reconstruction relative to a similar model in which the forward model was fixed. The simulation results showed that the SOFOMORE model was able to reduce large distance errors also over the widely used minimum norm model. A serious concern of extending source localization methods to include forward model reconstruction is data over-fitting. However, the choice of a flexible ARD prior on the *change* in the forward fields reduces over-fit by only allowing for reconstruction in forward fields where it is relevant.

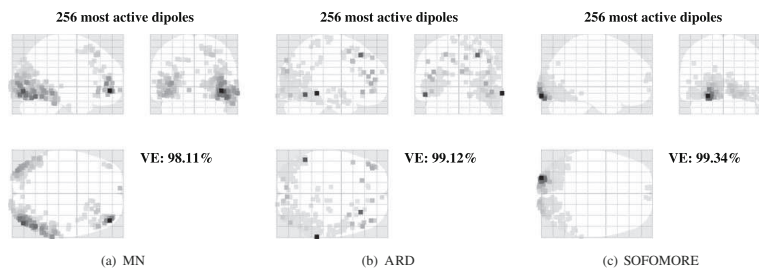


Figure 9 Estimated activity at $t = 170$ ms after stimulus. Tissue conductivities brain:skull:scalp = 0.33:0.0041:0.33 S/m are used. Activity in the left and right occipital region is estimated by MN with the primary activity located in the right occipital region. Moreover, right frontal activity is reconstructed. The ARD leads

to quite scattered activity with two dominating dipoles located in the left and right temporal lobe. SOFOMORE reconstructs activity both in the left and right visual cortex with dominating activity in the left region.

A.3 The $q(\alpha)$ Distribution

The variational posterior distribution for the hyperparameters of the sources is achieved by evaluating the Eq. 13 with $\mathbf{Z} = \alpha$. This leads to

$$q(\alpha) \propto \prod_{i=1}^{N_t} e^{\left\{ \frac{1}{2} \ln |\mathbf{D}| - \frac{1}{2} (\mathbf{s}_i^T \mathbf{D} \mathbf{s}_i) \right\} + \sum_{i=1}^{N_d} \{(v_{\alpha_i} - 1) \ln \alpha_i - \zeta_{\alpha_i} \alpha_i\}}$$

$$= \prod_{i=1}^{N_d} \alpha_i^{\left(v_{\alpha_i} + \frac{N_t}{2} - 1\right)} e^{-\left(\zeta_{\alpha_i} + \frac{1}{2} \sum_{i=1}^{N_t} (s_{it}^2)\right) \alpha_i} \quad (20)$$

from which we can infer that $q(\alpha)$ is a product of a set of N_d independent Gamma distributions and expressed as

$$q(\alpha) = \prod_{i=1}^{N_d} q(\alpha_i) = \prod_{i=1}^{N_d} \mathcal{G}(\alpha_i | \hat{v}_{\alpha_i}, \hat{\zeta}_{\alpha_i}). \quad (21)$$

Here the shape and rate parameters are given by

$$\hat{v}_{\alpha_i} = v_{\alpha_i} + \frac{N_t}{2}, \quad \hat{\zeta}_{\alpha_i} = \zeta_{\alpha_i} + \frac{1}{2} \sum_{i=1}^{N_t} (s_{it}^2), \quad (22)$$

respectively.

A.4 The $q(\beta)$ Distribution

In order to find the estimate of the variational distribution $q(\beta)$, we once more make use of Eq. 13 and now neglect all terms with functional independence on β ,

$$q(\beta) \propto e^{\left\{ \frac{N_t N_c}{2} \ln \beta - \frac{1}{2} \sum_{i=1}^{N_t} \langle (\mathbf{m}_i - \mathbf{A} \mathbf{s}_i)^2 \rangle + (v_\beta - 1) \ln \beta - \zeta_\beta \beta \right\}}$$

$$= \beta^{\left(v_\beta + \frac{N_t N_c}{2} - 1\right)} e^{-\left(\zeta_\beta + \frac{1}{2} \sum_{i=1}^{N_t} \langle (\mathbf{m}_i - \mathbf{A} \mathbf{s}_i)^2 \rangle\right) \beta}. \quad (23)$$

Based on Eq. 23, we see that the form of the distribution is a Gamma distribution

$$q(\beta) = \mathcal{G}(\beta | \hat{v}_\beta, \hat{\zeta}_\beta), \quad (24)$$

with shape and rate parameters

$$\hat{v}_\beta = v_\beta + \frac{N_t N_c}{2}, \quad \hat{\zeta}_\beta = \zeta_\beta + \frac{1}{2} \sum_{i=1}^{N_t} \langle (\mathbf{m}_i - \mathbf{A} \mathbf{s}_i)^2 \rangle. \quad (25)$$

A.5 The $q(\gamma)$ Distribution

The variational distribution $q(\gamma)$ is similarly obtained by neglecting all terms with no functional dependence on γ , which leads to

$$q(\gamma) \propto e^{\frac{N_t}{2} \ln |\mathbf{G}| - \frac{1}{2} \sum_{i=1}^{N_t} \left((\mathbf{I}_i - \mathbf{I}^{(0)})^T \mathbf{G} (\mathbf{I}_i - \mathbf{I}^{(0)}) \right)} \cdot e^{\sum_{i=1}^{N_t} \{(v_{\gamma_i} - 1) \ln \gamma_i - \zeta_{\gamma_i} \gamma_i\}}$$

$$= \prod_{i=1}^{N_d} \gamma_i^{\left(v_{\gamma_i} + \frac{N_t}{2} - 1\right)} e^{-\left(\zeta_{\gamma_i} + \frac{1}{2} \sum_{i=1}^{N_t} \left((\mathbf{a}_i - \mathbf{a}^{(0)})^2 \right)\right) \gamma_i} \quad (26)$$

From Eq. 26, we can infer that the $q(\gamma)$ distribution is given by the product of a set of Gamma distributions

$$q(\gamma) = \prod_{i=1}^{N_d} q(\gamma_i) = \prod_{i=1}^{N_d} \mathcal{G}(\gamma_i | \hat{v}_{\gamma_i}, \hat{\zeta}_{\gamma_i}), \quad (27)$$

where

$$\hat{v}_{\gamma_i} = v_{\gamma_i} + \frac{N_t}{2}, \quad \hat{\zeta}_{\gamma_i} = \zeta_{\gamma_i} + \frac{1}{2} \sum_{i=1}^{N_t} \left((\mathbf{a}_i - \mathbf{a}^{(0)})^2 \right). \quad (28)$$

A.6 Hyper-Hyperparameter Optimization

Since the model also includes a number of hyper-hyperparameters ($\{v_{\alpha_i}, \zeta_{\alpha_i}\}$, $\{v_\beta, \zeta_\beta\}$, and $\{v_{\gamma_i}, \zeta_{\gamma_i}\}$), these can also be optimized in order to increase the lower bound, and thereby make the lower bound as tight as possible to the log marginal likelihood. The hyper-hyperparameters can be regarded as parameter distributions with form of a Dirac delta function, which exactly will lead to point estimates of the hyper-hyperparameters [4]. In order to estimate the hyper-hyperparameters we evaluate the expected complete log likelihood w.r.t. the variational posterior distribution $q(\theta)$, and take the derivative w.r.t. the shape and rate parameter denoted v_x and ζ_x , respectively, for the parameter x . Thus, we have

$$\frac{\delta \mathcal{F}_{\mathcal{M}}}{\delta v_x} = 0 \Rightarrow \psi(v_x) = \ln \zeta_x - \langle \ln x \rangle \quad (29)$$

$$\frac{\delta \mathcal{F}_{\mathcal{M}}}{\delta \zeta_x} = 0 \Rightarrow \zeta_x = v_x / \langle x \rangle \quad (30)$$

where $\psi(v_x) = \delta \ln \Gamma(v_x) / \delta v_x$ is the digamma function and $\langle \ln x \rangle = \psi(\hat{v}_x) - \ln \hat{\zeta}_x$, [5]. Combining Eqs. 29 and 30 results in a nonlinear equation of form

$$f(v_x) = \psi(v_x) - \ln v_x + \ln \langle x \rangle - \langle \ln x \rangle = 0 \quad (31)$$

which we can solve using an iterative scheme like e.g., Newton–Raphson [27] and achieve an approximate solution of the real root that fulfills Eq. 31.

However, since the Gamma distribution is only defined for positive values of v_x and ζ_x , we must ensure that they also remain positive. This can be achieved with a formulation of Eq. 31 in the exponential domain, by performing the variable substitution, $v_x = e^\rho$, to obtain

$$f(\rho) = \psi(e^\rho) - \ln e^\rho + \ln(x) - \langle \ln x \rangle = 0. \quad (32)$$

By applying Newton–Raphsons methods, the update of ρ can be expressed as

$$\rho^{(k+1)} = \rho^{(k)} - \frac{f(\rho^{(k)})}{f'(\rho^{(k)})} \quad (33)$$

and with the use of $v_x = e^\rho$ and $\delta f(\rho)/\delta \rho = v_x \psi'(v_x) - 1$ we obtain the following update of the hyperhyperparameter v_x ,

$$v_x^{(k+1)} = v_x^{(k)} \exp \left[- \frac{\psi(v_x^{(k)}) - \ln v_x^{(k)} + \ln(x) - \langle \ln x \rangle}{v_x^{(k)} \psi'(v_x^{(k)}) - 1} \right] \quad (34)$$

Given an estimate of the v_x after performing the Newton–Raphson search, the shape parameter ζ_x is obtained using Eq. 30. The number of iterations performed in the Newton–Raphson search depends on the desired exactness. However, there is a trade-off between exactness and convergence time.

References

- Attias, H. (2000). A variational Bayesian framework for graphical models. *Advances in Neural Information Processing Systems*, 12(1–2), 209–215.
- Baillet, S., & Garnero, L. (1997). A Bayesian approach to introducing anatomo-functional priors in the EEG/MEG inverse problem. *IEEE Transactions on Biomedical Engineering*, 44(5), 374–385.
- Baillet, S., Mosher, J., & Leahy, R. (2001). Electromagnetic brain mapping. *IEEE Signal Processing Magazine*, 18, 14–30.
- Beal, M. (2003). *Variational algorithms for approximate Bayesian inference*. Ph.D. thesis, The Gatsby Computational Neuroscience Unit, University College London, UCL.
- Bishop, C. M. (2006). *Pattern recognition and machine learning*. New York: Springer.
- Daunizeau, J., Grova, C., Marrelec, G., Mattout, J., Jbabdi, S., Péligrini-Issac, M., et al. (2007). Symmetrical event-related EEG/fMRI information fusion in a variational bayesian framework. *NeuroImage*, 36, 69–87.

- Friston, K., Harrison, L., Daunizeau, J., Kiebel, S., Phillips, C., Trujillo-Barreto, N., et al. (2008). Multiple sparse priors for the M/EEG inverse problem. *NeuroImage*, 39, 1104–1120.
- Ghahramani, Z., & Beal, M. J. (2000). Graphical models and variational methods. In *Advanced mean field methods—Theory and practice*. MIT Press.
- Golub, G., Hansen, P., & O’Leary, D. (2000). Tikhonov regularization and total least squares. *SIAM Journal on Matrix Analysis and Applications*, 21(1), 185–194.
- Grova, C., Daunizeau, J., Lina, J.M., Bénar, C., Benali, H., & Gotman, J. (2006). Evaluation of EEG localization methods using realistic simulations of interictal spikes. *NeuroImage*, 29, 734–753.
- Hämäläinen, M., & Ilmoniemi, R. (1994). Interpreting magnetic fields of the brain: Minimum norm estimates. *Medical & Biological Engineering & Computing*, 32, 35–42.
- Henson, R., Goshen-Gottstein, Y., Ganel, T., Otten, L., Quayle, A., & Rugg, M. (2003). Electrophysiological and hemodynamic correlates of face perception, recognition and priming. *Cerebral Cortex*, 13, 793–805. <http://www.fil.ion.ucl.ac.uk/spm/doc/papers/cc-henson-2003.pdf>.
- Lew, S., Wolters, C., Anwander, A., Makeig, S., & MacLeod, R. (2007). Low resolution conductivity estimation to improve source localization. In *New frontiers in biomagnetism. Proc. of the 15th int. conf. on biomag. Int. congress series* (Vol. 1300, pp. 149–152).
- Mattout, J., Phillips, C., Penny, W., Rugg, M., & Friston, K. (2006). MEG source localization under multiple constraints: An extended Bayesian framework. *NeuroImage*, 30(3), 753–767.
- Mosher, J. C., Leahy, R. M., & Lewis, P. S. (1999). EEG and MEG: Forward solutions for inverse methods. *IEEE Transactions on Biomedical Engineering*, 46(3), 245–259.
- Mosher, J. C., Lewis, P. S., & Leahy, R. M. (1992). Multiple dipole modeling and localization from spatio-temporal MEG data. *IEEE Transactions on Biomedical Engineering*, 39(6), 541–557.
- Petersen, K., & Pedersen, M. (2007). *The matrix cookbook*. <http://matrixcookbook.com>. Version 20070905.
- Phillips, C., Rugg, M., & Friston, K. (2002). Anatomically informed basis functions for EEG source localisation: Combining functional and anatomical constraints. *NeuroImage*, 16(3), 678–695.
- Plis, S., George, J., Jun, S., Ranken, D., Volegov, P., & Schmidt, D. (2007). Probabilistic forward model for electroencephalography source analysis. *Physics in Medicine and Biology*, 52(17), 5309–5328.
- Sato, M. A., Yoshioka, T., Kajihara, S., Toyama, K., Goda, N., Doya, K., et al. (1999). Hierarchical bayesian estimation for MEG inverse problem. *NeuroImage*, 23(3), 806–826.
- Scherg, M., Bast, T., & Berg, P. (1999). Multiple source analysis of interictal spikes: Goals, requirements, and clinical value. *Journal of Clinical Neurophysiology*, 16, 214–224.
- Stahlhut, C., Mørup, M., Winther, O., & Hansen, L. K. (2009). Hierarchical Bayesian model for simultaneous EEG source and forward model reconstruction (SOFOMORE). In *Proceedings of the 2009 19th IEEE signal processing society workshop on machine learning for signal processing*.
- Stahlhut, C., Mørup, M., Winther, O., & Hansen, L. K. (2009). SOFOMORE: Combined EEG source and forward model reconstruction. In *6th IEEE int. symp. on biomed. imaging (ISBI): From nano to macro* (pp. 450–453).
- Trujillo-Barreto, N., Aubert-Vazquez, E., & Penny, W. (2008). Bayesian M/EEG source reconstruction with spatio-temporal priors. *NeuroImage*, 39, 318–335.

25. von Ellenrieder, N., Muravchik, C., & Nehorai, A. (2006). Effects of geometric head model perturbation on the EEG forward and inverse problems. *IEEE Transactions on Biomedical Engineering*, 53(3), 421–429.
26. Wolters, C., Anwander, A., Tricoche, X., Weinstein, D., Koch, M., & MacLeod, R. (2006). Influence of tissue conductivity anisotropy on EEG/MEG field and return current computation in a realistic head model: A simulation and visualization study using high-resolution finite element modeling. *Neuroimage*, 30(3), 813–826.
27. Ypma, T. (1995). Historical development of the Newton–Raphson method. *SIAM Review*, 37(4), 531–551.



Carsten Stahlhut received the B.Eng. and M.Sc. degree in Electrical Engineering in 2006 and 2008, respectively, from the Technical University of Denmark, Denmark. He is now pursuing a Ph.D. degree in Applied Mathematics at the Technical University of Denmark, Department of Informatics and Mathematical Modelling. His research interests include machine learning and biomedical data, with the focus on Bayesian inference for functional neuroimaging.



Morten Morup received his Ph.D. degree in Applied Mathematics at the Technical University of Denmark in 2008. He is currently Assistant Professor at the Cognitive Systems Group at

the Technical University of Denmark and a current member of the Technical Committee for the IEEE International Workshops on Machine Learning for Signal Processing (MLSP). His research interest is machine learning and biomedical data analysis and a major focus of his research has been dedicated to unsupervised learning with application to the analysis of neuroimaging data such as EEG and fMRI.



Ole Winther works in machine learning with applications in bioinformatics, data mining, 3G+ wireless communication and collaborative filtering. Ole Winther (M.Sc. Physics and Computer Science '94 and PhD Physics '98, both at University of Copenhagen (KU)) has published 50+ scientific papers and has previously held positions at Lund University and Center for Biological Sequence Analysis, Technical University of Denmark (DTU). He currently holds joint positions as associate professor at ISP, IMM, DTU and group leader at Bioinformatics, KU.



Lars Kai Hansen is a professor of digital signal processing at the Technical University of Denmark, Lyngby, where he also heads the Cognitive Systems Section. His research concerns adaptive signal processing and machine learning with applications in biomedicine and digital media. He has published more than 225 contributions on these subjects in journals, conferences, and books.

APPENDIX D

Sparse Spatio-temporal Inference of Electromagnetic Brain Sources

Carsten Stahlhut, Hagai T. Attias, David Wipf, Lars K. Hansen, and Srikan-
tan S Nagarajan. Sparse Spatio-temporal Inference of Electromagnetic Brain
Sources. *International Workshop on Machine Learning in Medical Imaging*.
September 2010.

Sparse Spatio-temporal Inference of Electromagnetic Brain Sources

Carsten Stahlhut^{1,*}, Hagai T. Attias², David Wipf³,
Lars K. Hansen¹, and Srikantan S. Nagarajan³

¹ DTU Informatics, Kgs. Lyngby, Denmark

² Golden Metallic Inc., San Francisco, CA, USA

³ University of California, San Francisco, CA, USA

cs@imm.dtu.dk, htattias@goldenmetallic.com,

dwipf@mrsc.ucsf.edu, lkh@imm.dtu.dk, sri@mrsc.ucsf.edu

Abstract. The electromagnetic brain activity measured via MEG (or EEG) can be interpreted as arising from a collection of current dipoles or sources located throughout the cortex. Because the number of candidate locations for these sources is much larger than the number of sensors, source reconstruction involves solving an inverse problem that is severely underdetermined. Bayesian graphical models provide a powerful means of incorporating prior assumptions that narrow the solution space and lead to tractable posterior distributions over the unknown sources given the observed data. In particular, this paper develops a hierarchical, spatio-temporal Bayesian model that accommodates the principled computation of sparse spatial and smooth temporal M/EEG source reconstructions consistent with neurophysiological assumptions in a variety of event-related imaging paradigms. The underlying methodology relies on the notion of automatic relevance determination (ARD) to express the unknown sources via a small collection of spatio-temporal basis functions. Experiments with several data sets provide evidence that the proposed model leads to improved source estimates. The underlying methodology is also well-suited for estimation problems that arise from other brain imaging modalities such as functional or diffusion weighted MRI.

Keywords: underdetermined inverse problems, M/EEG source reconstruction, probabilistic graphical models, variational Bayes, ARD.

1 Introduction

For both research and clinical purposes, a variety of imaging modalities have been applied to the human brain, including functional magnetic resonance (fMRI), positron emission tomography (PET), electro- and magneto-encephalography (EEG, MEG). While the models developed herein can ultimately benefit many imaging techniques, we focus our attention on EEG and MEG, both of which are very promising because of their excellent temporal resolution measuring neural

* Thanks: Support from Lundbeck Foundation through CIMBI – www.cimbi.org.

158 C. Stahlhut et al.

activity. The relation between the measured M/EEG signal and the brain's current sources can be expressed as a linear instantaneous form in the sources. The forward relation can be written as [1]

$$\mathbf{Y} = \mathbf{A}\mathbf{S} + \mathcal{E}, \quad (1)$$

where the noise \mathcal{E} is assumed additive and the measured M/EEG signal is denoted $\mathbf{Y} \in \mathbb{R}^{N_c \times N_t}$, the current sources $\mathbf{S} \in \mathbb{R}^{N_d \times N_t}$, and with N_c , N_d , and N_t being the number of channels, dipoles (or sources), and time samples, respectively. The coupling of sensors and the current sources is expressed through the lead field matrix/forward model $\mathbf{A} \in \mathbb{R}^{N_c \times N_d}$ with the rows referred to as the lead fields for the sensors and the columns as the forward fields for the sources. Multiple methods based on the physical properties of the brain and Maxwell's equations are available for computing \mathbf{A} .

To obtain reasonable spatial resolution, the number of source locations will necessarily be much larger than the number of sensors ($N_d \gg N_c$). Consequently, the inverse problem, estimating \mathbf{S} given \mathbf{Y} is severely underdetermined (ill-posed), and all efforts at source reconstruction are heavily dependent on prior assumptions, which may be viewed, either explicitly or implicitly, as placing some prior distribution on \mathbf{S} . Additionally, existing inverse algorithms face the challenge of the many sources of noise that interfere with the true signals in the M/EEG data. Electrical, thermal and biological noise as well as background room interference can be present. As a consequence of the many noisy contributions and the highly ill-posed nature of the electromagnetic source imaging (ESI) this leads to high requirements on robust inverse methods. In this paper we pursue a spatio-temporal method.

The use of both spatial and temporal information to constrain the source estimates has been advocated for decades. For example, temporal priors have been proposed that operate by penalizing differences in neighboring time points [1], while other priors incorporate temporal smoothness based on second derivatives [2]. To describe the temporal dynamics of the sources more flexibly, wavelet temporal basis functions have also received much attention in the M/EEG community. In [3] the focus is to represent event related potential (ERPs) with the use of a small set of wavelet bases. Additionally, [4] presents a variational Bayesian approach which tries to represent the M/EEG signal by a small set of coefficients using a wavelet shrinkage procedure.

In [5] an event sparse penalty (ESP) procedure is developed that seeks a solution composed of a small number of space-time events (STEs) chosen out of a large set of candidates (unlike the method from [4]). Each STE is a spatio-temporal signal defined in terms of a group of basis functions. Similarly, we pursue ESI solutions with sparse spatio-temporal representation. However, in contrast to [5], which only computes point estimates by solving a penalized generalized ℓ_1 norm regression problem, in Section 2 we develop a hierarchical Bayesian model, using ideas from probabilistic graphical modeling, which reflects uncertainty by allowing us to compute a tractable posterior distribution on the unknown sources. Later in Section 3 we provide empirical evidence of the efficacy of this approach.

2 Method

We propose a spatio-temporal model, which we will call Aquavit, for solving the ill-posed M/EEG source localization problem. The hierarchical structure of our Aquavit model is related to the models proposed in [4,6] and the Champagne algorithm from [7]. However, in contrast to [7], which makes the unrealistic assumption that a given current source is independent over time, we here restrict the current sources to be correlated by imposing a fixed temporal basis set $\Phi = [\phi_1, \dots, \phi_{N_k}]^T$ consisting of N_k basis functions of lengths N_t . A crucial aspect in general with spatio-temporal models that uses basis functions is the choice of the of the basis functions. There are many approaches that people have taken to form spatial and temporal basis functions, see e.g. [1,2,3,4,5,8,9]. Fixed basis functions include Fourier series, wavelet or discrete prolate spheroidal functions which specify bandwidth and temporal extent of the data. Learned basis functions include PCA, ICA and related approaches which learn basis functions from data. Similarly, we will assume the noise covariance matrix as fixed, since this as well can be learned to effectively suppress noisy factors (e.g., using a variational Bayesian factor model proposed in [10]). Given the linear relationship in (1) we define the current sources \mathbf{S} to encompass both temporal evolution and source noise. This is in contrast to the model in [6], which doesn't include the source noise term.

$$\mathbf{S} = \mathbf{W}^T \Phi + \mathbf{V} \quad (2)$$

with $\mathbf{W} = [\mathbf{w}_1, \dots, \mathbf{w}_{N_d}]$ being the spatio-temporal maps (also denoted the regression coefficients) that we are interested in finding and \mathbf{V} representing the source noise. In fact this latter term is one of the things that differentiate our model from most of the other spatio-temporal models. The combination of both temporal basis functions and the source noise term ensures that we are not only constrained to a subspace specified by the basis functions but also allows flexibility to search for source activity outside the span of the basis. More specifically our hierarchical model spells

$$p(\mathbf{Y} | \mathbf{S}) = \prod_n \mathcal{N}(\mathbf{y}_n | \mathbf{A} \mathbf{s}_n, \Sigma_\varepsilon) \quad (3)$$

$$p(\mathbf{S} | \mathbf{W}) = \prod_n \mathcal{N}(\mathbf{s}_n | \mathbf{W}^T \boldsymbol{\varphi}_n, \Gamma^{-1}) \quad (4)$$

$$p(\mathbf{W}) = \prod_i \mathcal{N}(\mathbf{w}_i | \mathbf{0}, \gamma_i^{-1} \mathbf{D}_i^{-1}) \quad (5)$$

where $\Gamma = \text{diag}(\gamma)$, $\mathbf{D}_i = \text{diag}(\alpha_i)$, Σ_ε represents the noise covariance at the sensors, and $\boldsymbol{\varphi}_n$ denotes all basis functions at sample n ($\boldsymbol{\varphi}_n = \Phi_n$). Note that the covariance of the mixing matrix \mathbf{W} is parametrized using γ and \mathbf{D}_i . We could have omitted γ from $p(\mathbf{W})$ without changing the model. However, the present parametrization simplifies the update rules. The prior distribution $p(\mathbf{W})$ enforces the notion of ARD at the level of both voxels and basis functions

160 C. Stahlhut et al.

such that if any element of α_i is large, the associated basis function at voxel i is deemed irrelevant to explaining the data. During the learning process (see below), many such elements are driven to infinity, effectively allowing the sources \mathbf{S} to be represented using only a modest set of basis functions. It is in this capacity that our model differs from [4].

Ideally, we would like to integrate out all unknowns (including hyperparameters) and then compute the posterior over sources $p(\mathbf{S}|\mathbf{Y})$, which contains all possible information about \mathbf{S} conditioned on the observed data \mathbf{Y} . However, the exact posterior in our model is computationally intractable. Here we choose to infer the hyperparameters using the maximum a-posteriori (MAP) point estimate. Given this, the joint posterior over the sources and mixing matrix $p(\mathbf{S}, \mathbf{W}|\mathbf{Y})$ is Gaussian, and can be computed analytically. Nevertheless, here we apply a variational Bayesian (VB) [11] approximation $p(\mathbf{S}, \mathbf{W}|\mathbf{Y}) \approx q(\mathbf{S}, \mathbf{W})$ to speed up the computation. In VB we maximize a lower bound of the log marginal likelihood

$$\begin{aligned} \mathcal{L} &= \log \int p(\mathbf{Y}, \mathbf{S}, \mathbf{W}) d\mathbf{S} d\mathbf{W} \geq \int q(\mathbf{S}, \mathbf{W}) \log \frac{p(\mathbf{Y}, \mathbf{S}, \mathbf{W})}{q(\mathbf{S}, \mathbf{W})} d\mathbf{S} d\mathbf{W} \\ &= \langle \log p(\mathbf{Y}, \mathbf{S}, \mathbf{W}) - \log q(\mathbf{S}, \mathbf{W}) \rangle_{q(\mathbf{S}, \mathbf{W})} = \mathcal{F} \end{aligned} \quad (6)$$

where $q(\mathbf{S}, \mathbf{W})$ is the joint variational distribution and $\langle \cdot \rangle$ denotes the expected value. Other than the factorial assumption $q(\mathbf{S}, \mathbf{W}) = q(\mathbf{S})q(\mathbf{W})$, we place no constraints on this distribution. We then iteratively maximize the objective function \mathcal{F} with respect to $q(\mathbf{S})$, $q(\mathbf{W})$, and any hyperparameters (i.e., $\mathbf{\Gamma}$, \mathbf{D}_i) we choose to update. Upon convergence, it can be shown that $q(\mathbf{S})$ becomes a principled Gaussian approximation to $p(\mathbf{S}|\mathbf{Y})$ with analytically computable moments.

Optimization of variational distributions: In turn each of the marginal posterior distributions can be maximized as follows.

$$q(\mathbf{S}) \propto \exp\left(\langle \log p(\mathbf{Y}, \mathbf{S}, \mathbf{W}) \rangle_{q(\mathbf{W})}\right) = \prod_n \mathcal{N}(\mathbf{s}_n | \boldsymbol{\mu}_n, \boldsymbol{\Sigma}_s) \quad (7)$$

$$\boldsymbol{\mu}_n = \boldsymbol{\Sigma}_s (\mathbf{A}^T \boldsymbol{\Sigma}_\varepsilon^{-1} \mathbf{y}_n + \mathbf{\Gamma} \langle \mathbf{W} \rangle^T \boldsymbol{\varphi}_n), \quad \boldsymbol{\Sigma}_s^{-1} = \mathbf{A}^T \boldsymbol{\Sigma}_\varepsilon^{-1} \mathbf{A} + \mathbf{\Gamma}. \quad (8)$$

Similarly, we maximize the variational posterior distribution of $q(\mathbf{W})$ by

$$q(\mathbf{W}) \propto \exp\left(\langle \log p(\mathbf{Y}, \mathbf{S}, \mathbf{W}) \rangle_{q(\mathbf{S})}\right) = \prod_i \mathcal{N}(\mathbf{w}_i | \boldsymbol{\eta}_i, \gamma_i^{-1} \boldsymbol{\Omega}_i) \quad (9)$$

$$\boldsymbol{\eta}_i = \boldsymbol{\Omega}_i \boldsymbol{\Phi} \langle \mathbf{S}_i^T \rangle, \quad \boldsymbol{\Omega}_i^{-1} = \boldsymbol{\Phi} \boldsymbol{\Phi}^T + \mathbf{D}_i. \quad (10)$$

Optimization of hyperparameters: In order for the variational posterior distributions to be optimal we need to maximize the hyperparameters as well. Following a VB expected maximization framework the hyperparameters is updated by taking the derivative of the expected complete data log likelihood w.r.t. the hyperparameters. However, the convergence can be extremely slow. Following Mackay [12], we derive fixed-point update rules, which in practice lead to

significantly faster convergence. Using the above variational posterior and some algebra, we obtain for the objective function \mathcal{F} , within a constant

$$\begin{aligned} 2\mathcal{F} = & N_t \log |\Sigma_\varepsilon^{-1}| + N_t \log |\Gamma| + \sum_i \log |\mathbf{D}_i| - N_t \log |\Sigma_s^{-1}| - \sum_i \log |\Omega_i| \\ & - \sum_n (\mathbf{y}_n - \mathbf{A} \langle \mathbf{s}_n \rangle)^T \Sigma_\varepsilon^{-1} (\mathbf{y}_n - \mathbf{A} \langle \mathbf{s}_n \rangle) - \sum_i \langle \mathbf{w}_i \rangle^T \gamma_i \mathbf{D}_i \langle \mathbf{w}_i \rangle \\ & - \sum_n (\langle \mathbf{s}_n \rangle - \langle \mathbf{W} \rangle^T \varphi_n)^T \Gamma (\langle \mathbf{s}_n \rangle - \langle \mathbf{W} \rangle^T \varphi_n). \end{aligned} \quad (11)$$

To derive the Mackay update for γ , we take the derivative

$$\begin{aligned} 2\partial\mathcal{F}/\partial\gamma_i = & N_t \gamma_i^{-1} - N_t (\Sigma_s)_{ii} - r_i = 0 \\ r_i = & \langle \mathbf{w}_i \rangle^T \mathbf{D}_i \langle \mathbf{w}_i \rangle + \sum_n (\langle \mathbf{s}_n \rangle - \langle \mathbf{W} \rangle^T \varphi_n)_i^2 \end{aligned} \quad (12)$$

and define a matrix \mathbf{H} which we will use to express Σ_s such that

$$\Sigma_s = \Gamma^{-1} (\mathbf{I} - \mathbf{H}) \Leftrightarrow \mathbf{H} = \mathbf{I} - \Gamma \Sigma_s = \mathbf{A}^T \Sigma_\varepsilon^{-1} \mathbf{A} \Sigma_s \Rightarrow \gamma_i = N_t \mathbf{H}_{ii} / r_i \quad (13)$$

in which we have substituted the expression for Σ_s in (12). Similarly, the update of the hyperparameter α_{ki} is given by

$$2\partial\mathcal{F}/\partial\alpha_{ki} = \alpha_{ki}^{-1} - (\Omega_i)_{kk} - c_{ki} = 0, \quad c_{ki} = \gamma_i \langle \mathbf{W} \rangle_{ki}^2 \quad (14)$$

where we can express Ω_i by a matrix \mathbf{G}_i and substitute such an expression into (14) in order to obtain the update of α_{ki}

$$\Omega_i = \mathbf{D}_i^{-1} (\mathbf{I} - \mathbf{G}_i) \Leftrightarrow \mathbf{G}_i = \mathbf{I} - \mathbf{D}_i \Omega_i = \Phi \Phi^T \Omega_i \Rightarrow \alpha_{ki} = (\mathbf{G}_i)_{kk} / c_{ki}. \quad (15)$$

In some situations, it is convenient to utilize a single α value to modulate groups of voxels or basis functions, in which case the associated update from (15) reduces to an average across the group.

3 Empirical Evaluation

In order to test the performance of the spatio-temporal Aquavit algorithm we use as benchmark two state-of-the-art methods; the Champagne algorithm [7] and the event sparse penalty (ESP) formulation in [5]. In Champagne a major difference from here is that the source time courses are random variables and need to be inferred from data, whereas here we have basis functions which are known in advance. We will denote the latter method as the ESP model. On the other hand the ESP is dependent on obtaining a set of basis functions that are very descriptive of the data.

Random lead field matrix: We first apply the models to a small toy problem where we construct a random generated lead field matrix ($\mathbf{A} : 10 \times 25$) with unit

162 C. Stahlhut et al.

variance. We generate two active sources with correlated times series, as seen in Fig. 1. Their times series are created by a sparse set of sine functions. More specifically, we have $s_1(t) = 0.3 \sin(4\pi t) + \sin(60\pi t)$ and $s_2(t) = \sin(10\pi t) + 0.8 \sin(60\pi t)$. Since noise is normally present in electromagnetic recordings, we apply a $\text{SNR} = 1$, with SNR defined as the ratio between the power of the clean signal and the noise.

Our set of temporal basis functions are constructed to include $N_k = 30$ sinusoids with frequencies between 1 and 30 Hz, denoted Φ^* , with * indicating that the temporal basis functions include the simulated sources. In contrast we construct a similar temporal basis set Φ , however, now with two of the sine functions (frequency 2 and 5 Hz) used to represent the simulated sources removed. By comparing the simulated sources in Fig. 1(a) and reconstructed sources in Fig. 1(b) it is seen that the Champagne algorithm performs quite well. Looking more carefully it is seen that the estimated temporal signatures are corrupted by a number of spikes, since this model reconstructs each sample independent of each other. In contrast the ESP model is very dependent on obtaining the correct temporal subspace. Given the temporal basis set Φ^* the ESP almost reconstructs the simulated sources perfectly, Fig. 1(c). However, given a more likely situation, where we cannot be sure that the temporal basis set include all the active sources time series, the source estimates obtained by the ESP model deteriorates quite significantly in Fig. 1(d). Since, ESP model is restricted to the pre-defined set of temporal basis functions, the model is not able to represent the simulated sources with time series outside the temporal basis set. Indeed, the interpretation that sources both consists of some temporal patterns and ‘source

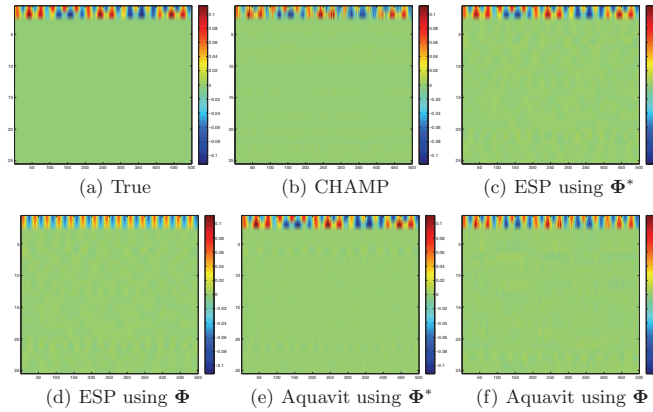


Fig. 1. Reconstructed source activity with y-axis expressing the index number of the sources and the x-axis representing the time. Mean square error (MSE) on the source estimates: (a) 0, (b) 0.179, (c) 0.011, (d) 0.408, (e) 0.015, and (f) 0.160.

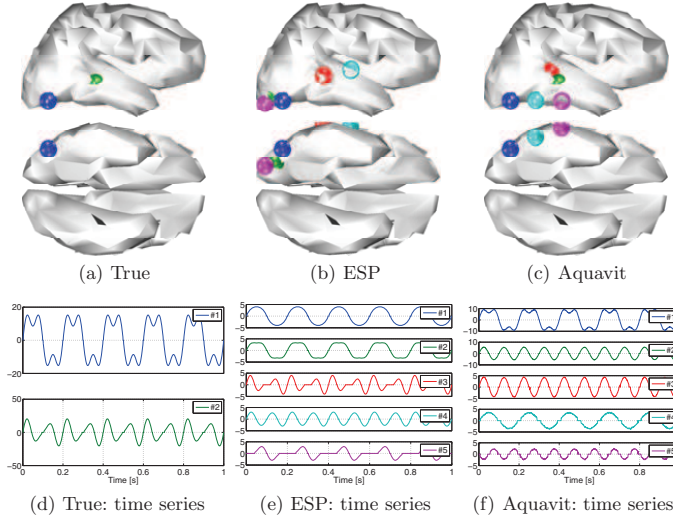


Fig. 2. Brain: The five sources with the largest variances. The sources are color-coded such that their corresponding time courses can be seen below the brain images. MSE/AUC measures: (a) 0/1, (b) 0.862/0.541, and (c) 0.403/1.

noise' actually leads to a balance between these two quantities. Given the Φ^* the temporal signature of the reconstructed sources Φ^* are almost perfectly recovered in Fig. 1(e). Note how the source noise quantity is taken over in Fig. 1(f) and obtains similar patterns as the Champagne algorithm.

Realistic Head Model: To illustrate the applicability of the Aquavit algorithm on more realistic problems we apply the model to a cortical resolution of (~ 500 vertices) obtained by subsampling a more dense set. However, such low resolution or even lower resolution might actually be the case if applied in a context similar to [5], in which vertices are clustered such that they form a number of patches based on anatomical and functional information. As head model we use a Boundary Element Method (BEM) forward model from SPM8¹. We compare the Aquavit algorithm with the ESP model. Similar to the toy problem we generate two simulated current sources from a mixture of two sines-functions with different frequencies. In the simulations we add random noise such that a SNR = 0.5 is obtained. Figure 2(a) and 2(d) show the locations and times series for the two simulated sources, respectively. In Fig. 2(b) the ESP model localizes the position of source #1 correctly but the temporal signature is not correctly recovered. The Aquavit algorithm on the other hand manages to reconstruct both

¹ SPM8 academic software (<http://www.fil.ion.ucl.ac.uk/spm/>).

164 C. Stahlhut et al.

position and time series of source #1. For source #2 the time series seem better recovered by the ESP model, however with the position of the reconstructed source being misplaced. This misplacement is also reflected in the area under the receiver operating curve (AUC) of 0.541. In contrast Aquavit results in AUC=1, which indicates perfect localization. However, minor deviation in the temporal signature of source #2 exists.

4 Discussion

This paper derives a hierarchical spatio-temporal model that seeks a sparse spatial and smooth temporal representation of the M/EEG using the principle of automatic relevance determination. Simulations demonstrated the potential of this spatio-temporal model with its ability to balance prior temporal information and ‘noise’ at the source level.

References

1. Baillet, S., Garnero, L.: A bayesian approach to introducing anatomo-functional priors in the EEG/MEG inverse problem. *IEEE Trans. on Biomed. Eng.* 44(5), 374–385 (1997)
2. Daunizeau, J., Mattout, J., Clonda, D., Goulard, B., Benali, H., Lina, J.M.: Bayesian spatio-temporal approach for EEG source reconstruction: Conciliating ECD and distributed models. *IEEE Trans. on Biomed. Eng.* 53(3), 503–516 (2006)
3. Effern, A., Lehnertz, K., Fernandez, G., Grunwald, T., David, P., Elger, C.E.: Single trial analysis of event related potentials: non-linear de-noising with wavelets. *Clinical neurophysiology* 111(12), 2255–2263 (2000)
4. Trujillo-Barreto, N.J., Aubert-Vazquez, E., Penny, W.D.: Bayesian M/EEG source reconstruction with spatio-temporal priors. *NeuroImage* 39, 318–335 (2008)
5. Bolstad, A., Van Veen, B., Nowak, R.: Spacetime event sparse penalization for magneto-/electroencephalography. *NeuroImage* 46(4), 1066–1081 (2009)
6. Sahani, M., Nagarajan, S.S.: Reconstructing meg sources with unknown correlations. In: *Advances in Neural Information Processing Systems*, vol. 16. MIT Press, Cambridge (2004)
7. Wipf, D.P., Owen, J.P., Attias, H.T., Sekihara, K., Nagarajan, S.S.: Robust Bayesian estimation of the location, orientation, and time course of multiple correlated neural sources using MEG. *Neuroimage* 49(1), 641–655 (2010)
8. Zumer, J.M., Attias, H.T., Sekihara, K., Nagarajan, S.S.: Probabilistic algorithms for meg/eeeg source reconstruction using temporal basis functions learned from data. *NeuroImage* 41, 924–940 (2008)
9. Dyrholm, M., Makeig, S., Hansen, L.K.: Model selection for convolutive ICA with an application to spatiotemporal analysis of EEG. *Neural computation* 19(4), 934–955 (2007)
10. Nagarajan, S.S., Attias, H.T., Hild, K.E., Sekihara, K.: A probabilistic algorithm for robust interference suppression in bioelectromagnetic sensor data. *Statistics in Medicine* 26(21), 3886–3910 (2007)
11. Attias, H.: A variational Bayesian framework for graphical models. *Advances in neural information processing systems* 12(1-2), 209–215 (2000)
12. MacKay, D.J.C.: Bayesian interpolation. *Neural computation* 4(3), 415–447 (1992)

APPENDIX E

Evaluation of the Influence of uncertain Forward Models on the EEG Source Reconstruction Problem

Carsten Stahlhut, Morten Mørup, Ole Winther, and Lars K. Hansen. Evaluation of the Influence of uncertain Forward Models on the EEG Source Reconstruction Problem. *NeuroImage*. Vol.47. Proceedings of the 15th Annual Human Brain Mapping Meeting. June 2009.

Evaluation of the Influence of uncertain Forward Models on the EEG Source Reconstruction Problem

Carsten Stahlhut, Morten Mørup, Ole Winther, and Lars Kai Hansen

Department of Informatics and Mathematical Modelling, Technical University of Denmark, Richard Petersens Plads, 2800 Kgs. Lyngby, Denmark.

Introduction

Electro-encephalography (EEG) holds great promise for functional brain imaging, due to its high temporal resolution, low cost equipment and the possibility of performing the experiments under much more realistic conditions as compared to functional magnetic resonance imaging and positron emission tomography. Today's EEG brain imaging methods operate with the assumption that the forward model is known when the source estimation is performed. Many sources of uncertainty are involved in the formulation of the forward model like tissue segmentation, tissue conductivities, and electrode locations. In this contribution we investigate how forward model uncertainty influences source localization.

Methods

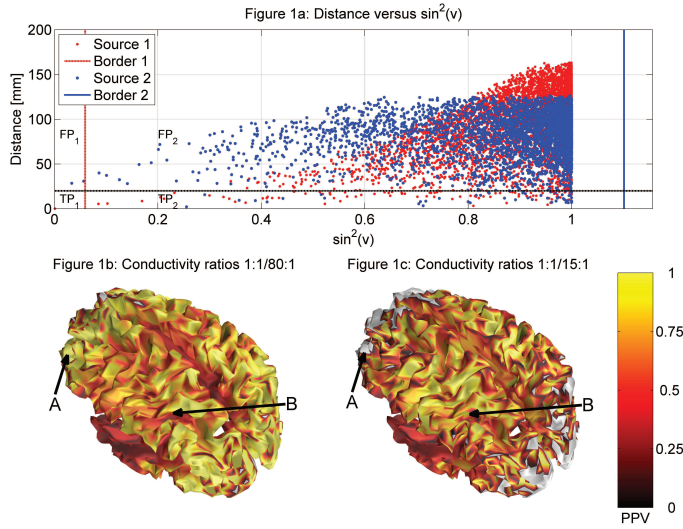
The analysis were based on 3-spheres models, where a high-resolution reference head model denoted as the 'true forward model' were compared with lower resolution forward models with and without erroneous tissue conductivity values. Conductivities brain:skull:scalp=0.33:0.0041:0.33S/m (ratio 1:1/80:1), were used in the true forward model and 1:1/15:1 in an erroneous model. To reveal the influence of the forward fields on the source estimates, we base our analysis on a simple 'stepwise' selection procedure, where a squared error function is used. For simplicity we assume that the true source configuration consists of a single dipole i and we now evaluate the cost estimate of a single dipole solution located at the site j . This allows us to examine how the dipoles are confused in the different areas of the brain when noise is present.

Results

Due to mismatch between the true and experimental forward model, the reconstruction of the sources is determined by the angles between the i 'th forward field associated with the true source and the j 'th forward field in the experimental forward model. Figure 1a shows two examples of confusion of the reconstructed sources when the true source is located in left frontal region (Source 1) and left temporal lobe (Source 2). The left side on the vertical lines indicates trusted regions where the cost of selecting one of the sources is smaller than the contribution from noise. As confusion measure we use: The positive prediction value $PPV = TP / (TP + FP)$, where TP is true positives (distance ≤ 20 mm) and FP is false positives (distance > 20 mm) both with angular factors smaller than the effective noise level. Figure 1b-1c show the PPV's for the whole brain with the true and erroneous conductivities, respectively. White areas indicate that no TP or FP has been detected. Generally, small signals from sulci and from cortical regions at a large distance from the sensors are more likely to be confused since the differences in angular factors can be small compared to the effective noise level. Increasing 'white' areas are found in figure 1c as a result of the poorer signal-to-noise ratio. Confusion is smaller for sources in the parietal region with the erroneous conductivity model, however, the angular factors also increases indicating a poorer representation of the signal.

Conclusion

This analysis demonstrated that caution is needed when evaluating the source estimates in different brain regions. Moreover, we demonstrated the importance of reliable forward models, which may be used as a motivation for including the forward model uncertainty into the source reconstruction methods.



References

- Homma, S., Musha, T., Nakjima, Y., Okamoto, Y., Blom, S., Flink, R., & Hagbarth, K.E. (1995), 'Conductivity ratios of the scalp-skull-brain head model in estimating equivalent dipole sources in human brain', *Neuroscience Research*, vol. 22, no. 1, pp. 51-55.
- Mosher, J.C., Leahy, R.M., & Lewis, P.S. (1999), 'EEG and MEG: forward solutions for inverse methods', *IEEE Transaction on Biomedical Engineering*, vol. 46, no. 3, pp. 245-259.
- Oostendorp, T. F., Delbeke, J., & Stegeman, D. F. (2000), 'The conductivity of the human skull: results of in vivo and in vitro measurements', *IEEE Transactions on Biomedical Engineering*, vol. 47, no. 11, pp. 1487-1492.

Bibliography

- [Almli et al., 2007] Almli, C., Rivkin, M., and McKinstry, R. (2007). The NIH MRI study of normal brain development (Objective-2): Newborns, infants, toddlers, and preschoolers. *Neuroimage*, 35(1):308–325.
- [Altman and Bernal, 2001] Altman, N. and Bernal, B. (2001). Brain Activation in Sedated Children: Auditory and Visual Functional MR Imaging1. *Radiology*, 221(1):56.
- [Ataseven et al., 2008] Ataseven, Y., Akalin-Acar, Z., Acar, C., and Gençer, N. (2008). Parallel implementation of the accelerated BEM approach for EMSI of the human brain. *Medical and Biological Engineering and Computing*, 46(7):671–679.
- [Attias, 2000] Attias, H. (2000). A variational Bayesian framework for graphical models. *Advances in neural information processing systems*, 12(1-2):209–215.
- [Auranen et al., 2005] Auranen, T., Nummenmaa, A., and Matti S. Hämäläinen, Iiro P. Jääskeläinen, J. L. A. V. M. S. (2005). Bayesian analysis of the neuromagnetic inverse problem with l^p -norm priors. *NeuroImage*, 26:870–884.
- [Baillet and Garnero, 1997] Baillet, S. and Garnero, L. (1997). A bayesian approach to introducing anatomo-functional priors in the EEG/MEG inverse problem. *IEEE Trans. on Biomedical Engineering*, 44(5):374–385.
- [Baillet et al., 2001] Baillet, S., Mosher, J., and Leahy, R. (2001). Electromagnetic brain mapping. *IEEE Signal Processing Magazine*, 18:14–30.

- [Baysal and Haueisen, 2004] Baysal, U. and Haueisen, J. (2004). Use of a priori information in estimating tissue resistivities-application to human data in vivo. *Physiological Measurement*, 25(3):737–748.
- [Beal, 2003] Beal, M. (2003). *Variational Algorithms for Approximate Bayesian Inference*. PhD thesis, The Gatsby Computational Neuroscience Unit, University College London, UCL.
- [Bentin et al., 1996] Bentin, S., Allison, T., Puce, A., Perez, E., and McCarthy, G. (1996). Electrophysiological studies of face perception in humans. *Journal of cognitive neuroscience*, 8(6):551–565.
- [Bernal et al., 2010] Bernal, S., Dehaene-Lambertz, G., Millotte, S., and Christophe, A. (2010). Two-year-olds compute syntactic structure on-line. *Developmental Science*, 13(1):69–76.
- [Bishop, 2006] Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. Springer, NY 10013 (USA).
- [Bolstad et al., 2009] Bolstad, A., Veen, B. V., and Nowak, R. (2009). Space-time event sparse penalization for magneto-/electroencephalography. *NeuroImage*, 46(4):1066–1081.
- [Bressler et al., 2007] Bressler, S., Richter, C., Chen, Y., and Ding, M. (2007). Cortical functional network organization from autoregressive modeling of local field potential oscillations. *Statistics in medicine*, 26(21):3875–3885.
- [Chen et al., 2010] Chen, F., Hallez, H., and Staelens, S. (2010). Influence of skull conductivity perturbations on EEG dipole source analysis. *Medical physics*, 37(8):4475–4484.
- [Cheour et al., 2004] Cheour, M., Imada, T., Taulu, S., Ahonen, A., Salonen, J., and Kuhl, P. (2004). Magnetoencephalography is feasible for infant assessment of auditory discrimination. *Experimental Neurology*, 190(Supplement 1):44 – 51. Neuroimaging of Functional Development in Fetuses and Newborns.
- [Daunizeau et al., 2006] Daunizeau, J., Mattout, J., Clonda, D., Goulard, B., Benali, H., and Lina, J. M. (2006). Bayesian spatio-temporal approach for EEG source reconstruction: Conciliating ECD and distributed models. *Biomedical Engineering, IEEE Transactions on*, 53(3):503–516.
- [Del Cul et al., 2007] Del Cul, A., Baillet, S., and Dehaene, S. (2007). Brain dynamics underlying the nonlinear threshold for access to consciousness. *PLoS Biol*, 5(10):e260.
- [Dyrholm et al., 2007] Dyrholm, M., Makeig, S., and Hansen, L. K. (2007). Model selection for convolutive ICA with an application to spatiotemporal analysis of EEG. *Neural computation*, 19(4):934–955.

- [Effern et al., 2000] Effern, A., Lehnertz, K., Fernandez, G., Grunwald, T., David, P., and Elger, C. (2000). Single trial analysis of event related potentials: non-linear de-noising with wavelets. *Clinical neurophysiology*, 111(12):2255–2263.
- [Fawcett, 2006] Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters*, 27(8):861–874.
- [Friston et al., 2008] Friston, K., Harrison, L., Daunizeau, J., Kiebel, S., Phillips, C., Trujillo-Barreto, N., Henson, R., Flandin, G., and Mattout, J. (2008). Multiple sparse priors for the M/EEG inverse problem. *NeuroImage*, 39:1104–1120.
- [Friston et al., 2007] Friston, K., Mattout, J., Trujillo-Barreto, N., Ashburner, J., and Penny, W. (2007). Variational free energy and the laplace approximation. *NeuroImage*, 34(1):220 – 234.
- [Fuchs et al., 1998] Fuchs, M., Drenckhahn, R., Wischmann, H., and Wagner, M. (1998). An improved boundary element method for realistic volume-conductor modeling. *IEEE Transactions on Biomedical Engineering*, 45(8):980–997.
- [Fuchs et al., 2002] Fuchs, M., Kastner, J., Wagner, M., Hawes, S., and Ebersole, J. (2002). A standardized boundary element method volume conductor model. *Clinical Neurophysiology*, 113(5):702–712.
- [Fuchs et al., 1999] Fuchs, M., Wagner, M., Kohler, T., and Wischman, H. (1999). Linear and nonlinear current density reconstructions. *J. Clin. Neurophysiol.*, 16(3):267–295.
- [Gelman et al., 2003] Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B. (2003). *Bayesian Data Analysis*. Chapman & Hall/CRC, second edition.
- [Gelskov and Kouider, 2010] Gelskov, S. and Kouider, S. (2010). Psychophysical thresholds of face visibility during infancy. *Cognition*, 114(2):285–292.
- [Gelskov, 2008] Gelskov, S. A. V. (2008). Exploring the behavioral and neural correlates of face visibility in preverbal infants. Master’s thesis, University of Copenhagen.
- [Gencer and Acar, 2004] Gencer, N. and Acar, C. (2004). Sensitivity of EEG and MEG measurements to tissue conductivity. *Physics in Medicine and Biology*, 49(5):701–717.
- [Ghahramani and Beal, 2000] Ghahramani, Z. and Beal, M. J. (2000). Graphical models and variational methods. In *Advanced Mean Field Methods - Theory and Practice*. MIT Press.

- [Golub et al., 2000] Golub, G. H., Hansen, P. C., and O’Leary, D. P. (2000). Tikhonov regularization and total least squares. *SIAM journal on matrix analysis and applications*, 21(1):185–194.
- [Goncalves et al., 2003] Goncalves, S., de Munck, J., Verbunt, J., Bijma, F., Heethaar, R., and Lopes da Silva, F. (2003). In vivo measurement of the brain and skull resistivities using an eit-based method and realistic models for the head. *IEEE Transactions on Biomedical Engineering*, 50(6):754–767.
- [Grova et al., 2006] Grova, C., Daunizeau, J., Lina, J. M., Bénar, C., Benali, H., and Gotman, J. (2006). Evaluation of EEG localization methods using realistic simulations of interictal spikes. *NeuroImage*, 29:734–753.
- [Hallez et al., 2007] Hallez, H., Vanrumste, B., Grech, R., Muscat, J., De Clercq, W., Vergult, A., D’Asseler, Y., Camilleri, K., Fabri, S., Van Huffel, S., et al. (2007). Review on solving the forward problem in EEG source analysis. *Journal of NeuroEngineering and Rehabilitation*, 4(1):46.
- [Hallez et al., 2005] Hallez, H., Vanrumste, B., Hese, P., D’Asseler, Y., Lemahieu, I., and Walle, R. (2005). A finite difference method with reciprocity used to incorporate anisotropy in electroencephalogram dipole source localization. *Physics in medicine and biology*, 50:3787–3806.
- [Hämäläinen and Ilmoniemi, 1994] Hämäläinen, M. and Ilmoniemi, R. (1994). Interpreting magnetic fields of the brain: minimum norm estimates. *Med. Biol. Eng. Comput.*, 32:35–42.
- [Hansen, 2010] Hansen, P. (2010). *Discrete inverse problems: insight and algorithms*. Society for Industrial & Applied.
- [Hansen, 1992] Hansen, P. C. (1992). Analysis of discrete ill-posed problems by means of the l-curve. *SIAM Review*, 34(4):pp. 561–580.
- [Haueisen et al., 1997] Haueisen, J., Ramon, C., Eiselt, M., Brauer, H., and Nowak, H. (1997). Influence of tissue resistivities on neuromagnetic fields and electric potentials studied with a finite element model of the head. *Biomedical Engineering, IEEE Transactions on*, 44(8):727–735.
- [Haufe et al., 2011] Haufe, S., Tomioka, R., Dickhaus, T., Sannelli, C., Blankertz, B., Nolte, G., and Müller, K.-R. (2011). Large-scale eeg/meg source localization with spatial flexibility. *NeuroImage*, 54(2):851 – 859.
- [Haufe et al., 2009] Haufe, S., Tomioka, R., Nolte, G., Mueller, K., and Kawanabe, M. (2009). Modeling sparse connectivity between underlying brain sources for EEG/MEG. *Arxiv preprint arXiv:0912.2412*.

- [Henson et al., 2003] Henson, R., Goshen-Gottstein, Y., Ganel, T., Otten, L., Quayle, A., and Rugg, M. (2003). Electrophysiological and hemodynamic correlates of face perception, recognition and priming. *Cerebral Cortex*, 13:793–805.
- [Henson et al., 2009] Henson, R., Mattout, J., Phillips, C., and Friston, K. (2009). Selecting forward models for meg source-reconstruction using model-evidence. *NeuroImage*, 46(1):168–176.
- [Homma et al., 1995] Homma, S., Musha, T., Nakajima, Y., Okamoto, Y., Blom, S., Flink, R., and Hagbarth, K.-E. (1995). Conductivity ratios of the scalp-skull-brain head model in estimating equivalent dipole sources in human brain. *Neuroscience Research*, 22(1):51–55.
- [Huiskamp et al., 1999] Huiskamp, G., Vroeijenstijn, M., van Dijk, R., Wieneke, G., and van Huffelen, A. (1999). The need for correct realistic geometry in the inverse EEG problem. *IEEE Trans. Biomed. Eng.*, 46(11):1281–1287.
- [Jacobsen et al., 2008] Jacobsen, D., Hansen, L., and Madsen, K. (2008). Bayesian model comparison in nonlinear bold fmri hemodynamics. *Neural computation*, 20(3):738–755.
- [Johnson et al., 1991] Johnson, M., Dziurawiec, S., Ellis, H., and Morton, J. (1991). Newborns’ preferential tracking of face-like stimuli and its subsequent decline. *Cognition*, 40(1-2):1–19.
- [Kouider et al., 2011] Kouider, S., Gelskov, S., Christophe, A., de Gardelle, V., Dehaene-Lambertz, G., and Stahlhut, C. (2011). Neural correlates of face visibility in infants. *in preparation*.
- [Lew et al., 2007] Lew, S., Wolters, C., Anwander, A., Makeig, S., and MacLeod, R. (2007). Low resolution conductivity estimation to improve source localization. In *New Frontiers in Biomagnetism. Proc. of the 15th Int. Conf. on Biomag.*, volume 1300 of *Int. Congress Series*, pages 149–152.
- [Liu et al., 2008] Liu, W.-C., Flax, J. F., Guise, K. G., Sukul, V., and Benasich, A. A. (2008). Functional connectivity of the sensorimotor area in naturally sleeping infants. *Brain Research*, 1223:42 – 49.
- [MacKay, 1992] MacKay, D. (1992). Bayesian interpolation. *Neural computation*, 4(3):415–447.
- [Minka, 2001] Minka, T. (2001). Expectation propagation for approximate Bayesian inference. In *Uncertainty in Artificial Intelligence*, volume 17, pages 362–369. Citeseer.

- [Morita et al., 2000] Morita, T., Kochiyama, T., Yamada, H., Konishi, Y., Yonekura, Y., Matsumura, M., and Sadato, N. (2000). Difference in the metabolic response to photic stimulation of the lateral geniculate nucleus and the primary visual cortex of infants: a fmri study. *Neuroscience Research*, 38(1):63 – 70.
- [Mørup, 2011] Mørup, M. (2011). Applications of tensor (multiway array) factorizations and decompositions in data mining. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 1(1):24–40.
- [Mosher et al., 1999] Mosher, J. C., Leahy, R. M., and Lewis, P. S. (1999). EEG and MEG: forward solutions for inverse methods. *IEEE Transactions on Biomedical Engineering*, 46(3):245–259.
- [Mosher et al., 1992] Mosher, J. C., Lewis, P. S., and Leahy, R. M. (1992). Multiple dipole modeling and localization from spatio-temporal MEG data. *IEEE Trans. on Biomedical Engineering*, 39(6):541–557.
- [Nagarajan et al., 2007] Nagarajan, S. S., Attias, H. T., Hild, K. E., and Sekihara, K. (2007). A probabilistic algorithm for robust interference suppression in bioelectromagnetic sensor data. *Statistics in Medicine*, 26(21):3886–3910.
- [Neal, 1997] Neal, R. (1997). Monte Carlo implementation of Gaussian process models for Bayesian regression and classification. *Arxiv preprint physics/9701026*.
- [Nunez and Srinivasan, 2006] Nunez, P. and Srinivasan, R. (2006). *Electric Fields of the Brain: The Neurophysics of EEG*. Oxford University Press, second edition.
- [Oostendorp et al., 2000] Oostendorp, T. F., Delbeke, J., and Stegeman, D. F. (2000). The conductivity of the human skull: results of in vivo and in vitro measurements. *IEEE Transactions on Biomedical Engineering*, 47(11):1487–1492.
- [Oostenveld and Oostendorp, 2002] Oostenveld, R. and Oostendorp, T. F. (2002). Validating the boundary element method for forward and inverse eeg computations in the presence of a hole in the skull. *Human Brain Mapping*, 17(3):179–192.
- [Oppen et al., 2008] Oppen, M., Paquet, U., and Winther, O. (2008). Improving on expectation propagation. *Advances in Neural Information Processing Systems*, 21:1241–1248.
- [Ou et al., 2009] Ou, W., Hämläinen, M., and Golland, P. (2009). A distributed spatio-temporal EEG/MEG inverse solver. *NeuroImage*, 44(3):932–946.

- [Pascual-Marqui et al., 1994] Pascual-Marqui, R. D., Michel, C. M., and Lehmann, D. (1994). Low resolution electromagnetic tomography: a new method for localizing electrical activity in the brain. *Int J. Psychophysiol.*, 18:49–65.
- [Patterson and Thompson, 1971] Patterson, H. D. and Thompson, R. (1971). Recovery of inter-block information when block sizes are unequal. *Biometrika*, 58(3):pp. 545–554.
- [Phillips, 2001] Phillips, C. (2001). *Source Estimation in EEG: Combining anatomical and functional constraints*. PhD thesis, Faculté des Sciences Appliquées, Université de Liège-Liège.
- [Phillips et al., 2002] Phillips, C., Rugg, M., and Friston, K. (2002). Anatomically Informed Basis Functions for EEG Source Localisation: Combining Functional and Anatomical Constraints. *NeuroImage*, 16(3):678–695.
- [Plis et al., 2007] Plis, S., George, J., Jun, S., Ranken, D., Volegov, P., and Schmidt, D. (2007). Probabilistic forward model for electroencephalography source analysis. *Physics in Medicine and Biology*, 52(17):5309–5328.
- [Plonsey and Heppner, 1967] Plonsey, R. and Heppner, D. (1967). Considerations of quasi-stationarity in electrophysiological systems. *Bulletin of Mathematical Biology*, 29:657–664. 10.1007/BF02476917.
- [Ramon et al., 2006] Ramon, C., Schimpf, P., and Haueisen, J. (2006). Influence of head models on EEG simulations and inverse source localizations. *Biomed. Eng. Online*, 5(10):1–13.
- [Reynolds et al., 2010] Reynolds, G., Courage, M., and Richards, J. (2010). Infant attention and visual preferences: Converging evidence from behavior, event-related potentials, and cortical source localization. *Developmental Psychology*, 46(4):886–904.
- [Reynolds and Richards, 2009] Reynolds, G. and Richards, J. (2009). Cortical source localization of infant cognition. *Developmental neuropsychology*, 34(3):312–329.
- [Richards, 2010a] Richards, J. E. (2010a). *Neoconstructivism: The new science of cognitive development*, chapter Attention in the brain and early infancy, pages 3–37. Oxford University Press, USA.
- [Richards, 2010b] Richards, J. E. (2010b). What’s inside a baby’s head? Structural and functional brain development in infants. In *17th Biennial International Conference on Infant Studies*.
- [Rush and Driscoll, 1969] Rush, S. and Driscoll, D. A. (1969). EEG electrode sensitivity- an application of reciprocity. *IEEE Transaction on Biomed. Eng.*, 16:15–22.

- [Sahani and Nagarajan, 2004] Sahani, M. and Nagarajan, S. S. (2004). Reconstructing meg sources with unknown correlations. In *Advances in Neural Information Processing Systems 16*. MIT Press, Cambridge, MA.
- [Saleheen and Ng, 1997] Saleheen, H. and Ng, K. (1997). New finite difference formulations for general inhomogeneous anisotropic bioelectric problems. *Biomedical Engineering, IEEE Transactions on*, 44(9):800–809.
- [Sato et al., 2004] Sato, M.-A., Yoshioka, T., Kajihara, S., Toyama, K., Goda, N., Doya, K., and Kawato, M. (2004). Hierarchical bayesian estimation for MEG inverse problem. *Neuroimage*, 23(3):806–826.
- [Scherg, 1990] Scherg, M. (1990). Fundamentals of dipole source potential analysis. *Auditory evoked magnetic fields and potentials*, 6:40–69.
- [Scherg et al., 1999] Scherg, M., Bast, T., and Berg, P. (1999). Multiple source analysis of interictal spikes: Goals, requirements, and clinical value. *J. Clin. Neurophysiol.*, 16:214–224.
- [Stahlhut et al., 2010a] Stahlhut, C., Attias, H. T., Wipf, D., Hansen, L. K., and Nagarajan, S. S. (2010a). Sparse spatio-temporal inference of electromagnetic brain sources. In Wang, F., Yan, P., Suzuki, K., and Shen, D., editors, *Machine Learning in Medical Imaging*, volume 6357 of *Lecture Notes in Computer Science*, pages 157–164. Springer Berlin / Heidelberg.
- [Stahlhut et al., 2011] Stahlhut, C., Gelskov, S., Dehaene-Lambertz, G., Hansen, L. K., and Kouider, S. (2011). On forward model selection in infants. *in preparation*.
- [Stahlhut et al., 2009a] Stahlhut, C., Mørup, M., Winther, O., and Hansen, L. K. (2009a). Hierarchical Bayesian model for simultaneous EEG source and forward model reconstruction (SOFOMORE). In *Machine Learning for Signal Processing, 2009. MLSP 2009. IEEE International Workshop on*, pages 1–6.
- [Stahlhut et al., 2009b] Stahlhut, C., Mørup, M., Winther, O., and Hansen, L. K. (2009b). SOFOMORE: Combined EEG source and forward model reconstruction. In *Biomedical Imaging: From Nano to Macro, 2009. ISBI '09. IEEE International Symposium on*, pages 450–453.
- [Stahlhut et al., 2010b] Stahlhut, C., Mørup, M., Winther, O., and Hansen, L. K. (2010b). Simultaneous eeg source and forward model reconstruction (sofomore) using a hierarchical bayesian approach. *Journal of Signal Processing Systems*, pages 1–14. 10.1007/s11265-010-0527-0.
- [Steinsträter et al., 2010] Steinsträter, O., Sillekens, S., Junghoefer, M., Burger, M., and Wolters, C. (2010). Sensitivity of beamformer source analysis to deficiencies in forward modeling. *Human Brain Mapping*, 31:1907–1927.

- [Strother et al., 2002] Strother, S., Anderson, J., Hansen, L., Kjems, U., Kustra, R., Sidtis, J., Frutiger, S., Muley, S., LaConte, S., and Rottenberg, D. (2002). The quantitative evaluation of functional neuroimaging experiments: The NPAIRS data analysis framework. *Neuroimage*, 15(4):747–771.
- [Tikhonov and Arsenin, 1977] Tikhonov, A. N. and Arsenin, V. Y. (1977). *Solutions of ill-posed problems*. John Wiley, New York.
- [Trujillo-Barreto et al., 2008] Trujillo-Barreto, N., Aubert-Vazquez, E., and Penny, W. (2008). Bayesian M/EEG source reconstruction with spatio-temporal priors. *NeuroImage*, 39:318–335.
- [Trujillo-Barreto et al., 2004] Trujillo-Barreto, N., Aubert-Vazquez, E., and Valdes-Sosa, P. (2004). Bayesian model averaging in EEG/MEG imaging. *NeuroImage*, 21(4):1300–1319.
- [Uutela et al., 1999] Uutela, K., Hämäläinen, M., and Somersalo, E. (1999). Visualization of magnetoencephalographic data using minimum current estimates. *NeuroImage*, 10:173–180.
- [Valdés-Sosa et al., 2000] Valdés-Sosa, P., Marti, F., Garcia, G., and Casanova, R. (2000). Variable resolution electric-magnetic tomography. *Biomag 96': Proceedings of the Tenth International Conference on Biomagnetism*, Vol. II, Springer-Verlag, New York:373–376.
- [Vanrumste et al., 2000] Vanrumste, B., Van Hoey, G., Van de Walle, R., D'Havé, M., Lemahieu, I., and Boon, P. (2000). Dipole location errors in electroencephalogram source analysis due to volume conductor model errors. *Medical and Biological Engineering and Computing*, 38:528–534. 10.1007/BF02345748.
- [Veen et al., 1997] Veen, B. D. V., van Drongelen, W., Yuchtman, M., and Suzuki, A. (1997). Localization of brain electrical activity via linearly constrained minimum variance spatial filtering. *Biomedical Engineering, IEEE Transactions on*, 44:867–880.
- [von Ellenrieder et al., 2006] von Ellenrieder, N., Muravchik, C., and Nehorai, A. (2006). Effects of geometric head model perturbation on the EEG forward and inverse problems. *IEEE Trans. Biomed. Eng.*, 53(3):421–429.
- [Wipf et al., 2010] Wipf, D. P., Owen, J. P., Attias, H. T., Sekihara, K., and Nagarajan, S. S. (2010). Robust Bayesian estimation of the location, orientation, and time course of multiple correlated neural sources using MEG. *Neuroimage*, 49(1):641–655.
- [Wolters et al., 2006] Wolters, C., Anwander, A., Tricoche, X., Weinstein, D., Koch, M., and MacLeod, R. (2006). Influence of tissue conductivity

- anisotropy on EEG/MEG field and return current computation in a realistic head model: A simulation and visualization study using high-resolution finite element modeling. *Neuroimage*, 30(3):813–826.
- [Wolters et al., 2004] Wolters, C., Grasedyck, L., and Hackbusch, W. (2004). Efficient computation of lead field bases and influence matrix for the FEM-based EEG and MEG inverse problem. *Inverse problems*, 20(4):1099–1116.
- [Wolters et al., 2007] Wolters, C., Kostler, H., Moller, C., Hardtlein, J., Grasedyck, L., and Hackbusch, W. (2007). Numerical mathematics of the subtraction method for the modeling of a current dipole in EEG source reconstruction using finite element head models. *SIAM J. Sci. Comp.*, 30(1):24–45.
- [Wolters et al., 2002] Wolters, C., Kuhn, M., Anwander, A., and Reitzinger, S. (2002). A parallel algebraic multigrid solver for finite element method based source localization in the human brain. *Computing and visualization in science*, 5(3):165–177.
- [Zumer et al., 2008] Zumer, J. M., Attias, H. T., Sekihara, K., and Nagarajan, S. S. (2008). Probabilistic algorithms for meg/eeg source reconstruction using temporal basis functions learned from data. *NeuroImage*, 41:924–940.